



Spécificités de l'implant électro-acoustique : indications, interface bioélectrique et stratégie de codage

Fabien Seldran

► To cite this version:

Fabien Seldran. Spécificités de l'implant électro-acoustique : indications, interface bioélectrique et stratégie de codage. Médecine humaine et pathologie. Université Claude Bernard - Lyon I, 2011. Français. NNT : 2011LYO10290 . tel-00751913

HAL Id: tel-00751913

<https://theses.hal.science/tel-00751913>

Submitted on 14 Nov 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N° d'ordre : 290-2011

Année 2011

THESE DE L'UNIVERSITE DE LYON

Délivrée par

L'UNIVERSITE CLAUDE BERNARD LYON 1

ECOLE DOCTORALE

Electronique, Electrotechnique, Automatique (EEA)

DISCIPLINE : Ingénierie pour le vivant

DIPLOME DE DOCTORAT

(arrêté du 7 août 2006)

soutenue publiquement le **19/12/2011**

par

M. Fabien SELDRAN

**SPÉCIFICITÉS DE L'IMPLANT ÉLECTRO-ACOUSTIQUE :
INDICATIONS, INTERFACE BIOÉLECTRIQUE ET STRATÉGIE DE CODAGE**

Directeur de thèse :
Pr. Christian BERGER-VACHON

Co-Directeur de thèse :
Pr. Hung THAI-VAN

JURY : Pr. Paul AVAN, **président**
M. Michel BELIAEFF
Pr. Christian BERGER-VACHON
Pr. Bruno FRACHET, **rapporteur**
Pr. Jean-Luc SCHWARTZ, **rapporteur**
Pr. Hung THAI-VAN
Pr. Eric TRUY

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université

M. A. Bonmartin

Vice-président du Conseil d'Administration

M. le Professeur G. Annat

Vice-président du Conseil des Etudes et de la Vie Universitaire

M. le Professeur D. Simon

Vice-président du Conseil Scientifique

M. le Professeur J-F. Mornex

Secrétaire Général

M. G. Gay

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard

Directeur : M. le Professeur J. Etienne

Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux

Directeur : M. le Professeur F-N. Gilly

UFR d'Odontologie

Directeur : M. le Professeur D. Bourgeois

Institut des Sciences Pharmaceutiques et Biologiques

Directeur : M. le Professeur F. Locher

Institut des Sciences et Techniques de la Réadaptation

Directeur : M. le Professeur Y. Matillon

Département de formation et Centre de Recherche en Biologie Humaine

Directeur : M. le Professeur P. Farge

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies

Directeur : M. le Professeur F. De Marchi

Département Biologie

Directeur : M. le Professeur F. Fleury

Département Chimie Biochimie

Directeur : Mme le Professeur H. Parrot

Département GEP

Directeur : M. N. Siauve

Département Informatique

Directeur : M. le Professeur S. Akkouche

Département Mathématiques

Directeur : M. le Professeur A. Goldman

Département Mécanique

Directeur : M. le Professeur H. Ben Hadid

Département Physique

Directeur : Mme S. Fleck

Département Sciences de la Terre

Directeur : Mme le Professeur I. Daniel

UFR Sciences et Techniques des Activités Physiques et Sportives

Directeur : M. C. Collignon

Observatoire de Lyon

Directeur : M. B. Guiderdoni

Ecole Polytechnique Universitaire de Lyon 1

Directeur : M. P. Fournier

Ecole Supérieure de Chimie Physique Electronique

Directeur : M. G. Pignault

Institut Universitaire de Technologie de Lyon 1

Directeur : M. le Professeur C. Coulet

Institut de Science Financière et d'Assurances

Directeur : M. le Professeur J-C. Augros

Institut Universitaire de Formation des Maîtres

Directeur : M. R. Bernard

REMERCIEMENTS

Je tiens avant tout à remercier le Pr Bruno Frachet, le Pr Jean-Luc Schwartz et le Pr Paul Avan qui me font l'honneur de participer à mon jury de thèse et de juger mon travail.

J'adresse mes plus sincères remerciements au Pr Christian Berger-Vachon, mon directeur de thèse, pour son encadrement, sa disponibilité et ses précieux conseils. Voilà bientôt 10 ans que je le connais, et durant ces 10 ans il m'a toujours soutenu et encouragé. A la fin de mon stage de licence au pavillon U, je suis allé le voir en lui disant « je souhaite continuer mon parcours universitaire dans l'audition ou dans les implants cochléaires », il m'a répondu qu'il m'aiderait du mieux qu'il pourrait et il a tenu promesse ! Est-ce la solidarité dauphinoise ? En tout cas, c'est grâce à lui si j'ai pu intégrer le domaine de l'audition et je lui en suis vivement reconnaissant.

J'aimerais également remercier le Pr Hung Thai-Van, mon co-directeur de thèse, pour son encadrement et sa collaboration, ainsi que pour les moyens qu'il a mis à ma disposition pour me permettre de mener au mieux mes recherches.

Merci au Pr Eric Truy qui m'accorde toute sa confiance. Sans lui je n'aurais probablement pas eu aussi facilement accès aux nombreux patients qui ont participé à mes études. Malgré un emploi du temps chargé, il a toujours trouvé un moment à m'accorder lorsque j'avais des questions, ou même lorsque j'avais besoin de discuter. La collaboration avec lui a été très enrichissante et je souhaite que nous puissions la prolonger après ma thèse.

Un grand MERCI à Michel Beliaeff, directeur de la société Vibrant Med-El France pour le financement de ma bourse CIFRE, ainsi que pour le support matériel et humain. Sans l'aide de Med-El, ce projet ne se serait certainement jamais réalisé. Je remercie également tous les autres membres de l'équipe qui m'accueillent toujours avec le sourire et qui sont toujours là pour moi, à savoir Samia Labassi, Bénédicte Philibert, Sandra Straube-Lassalle, Gregory Jacquot, Damien Dallais, Aymen Nsira, Marie Duroy, Agnès Pauvert. C'est avec un grand plaisir que je rejoins aujourd'hui la « famille » Med-El.

J'adresse également mes remerciements au Pr Lionel Collet qui m'a ouvert les portes de son service, ce qui m'a permis notamment de m'initier au réglage des implants cochléaires. Il m'a donné l'opportunité de suivre des études d'audioprothèse en parallèle de cette thèse et il m'a donné tous les moyens pour acquérir cette formation vraiment complète. J'ai particulièrement apprécié ses précieux conseils et le fait qu'il garde un œil sur mes travaux, malgré son mandat en tant que Président de l'Université Lyon 1 qui lui demandait beaucoup de temps et d'énergie. Pour donner un exemple, j'ai souvenir d'une réunion à Mandelieu entre Stéphane Gallégo, lui et moi, pendant laquelle il m'a fait la remarque « Et si tu prenais en compte la durée de privation pour analyser tes données ? ». Et 3 ans plus tard, ça nous a conduit à un JARO ! Pour tout cela je tiens à lui exprimer ma plus profonde reconnaissance.

Je tiens à remercier tout particulièrement le Dr Stéphane Gallégo pour son encadrement, son soutien, son investissement et ses précieux conseils sur le plan scientifique mais également sur le plan professionnel et personnel. Son enseignement et son expérience m'ont beaucoup appris. Le travail avec lui se fait, certes dans le stress, mais toujours dans la joie et la bonne humeur et c'est ce qui me permet d'avancer. C'est sans doute la personne envers laquelle je me sens la plus redevable tant il m'a apporté ces 5 dernières années et je ne sais d'ailleurs même pas comment le remercier. Il mériterait que je consacre 2 pages de remerciements rien que pour lui. Il a toujours été présent pour moi, à toute heure du jour et de la nuit, les week-end et jours fériés, parfois pour prendre ma défense ou pour me sortir du pétrin, et bien souvent pour répondre à mes innombrables « questions à 20 centimes ». J'aimerais aussi le remercier pour toutes les personnes qu'il m'a fait connaître ou avec qui il m'a mis en relation. Stéphane Gallégo est vraiment quelqu'un devant qui je suis admiratif et aujourd'hui c'est un honneur et une fierté de pouvoir dire que j'ai collaboré avec lui. Il aura été bien plus qu'un encadrant, c'est désormais un pote. C'est pour moi un modèle et j'espère vraiment suivre ses traces.

Je tiens également à remercier le Dr Christophe Micheyl, sans qui mes travaux n'auraient pas abouti et mes articles ne seraient probablement pas publiés aujourd'hui. J'ai fait sa connaissance en septembre 2010 et depuis un an et demi que nous travaillons ensemble, il a su valoriser mon travail et le faire fructifier sans commune mesure. Jamais je n'aurais pensé prétendre à un JARO en premier article et pourtant grâce à lui j'y suis parvenu. Son nom parmi les auteurs m'a peut-être aidé ? En tout cas c'est une fierté d'avoir mon nom à

côté du sien ! J'admire chez lui sa puissance de travail, sa détermination et par-dessus tout son esprit, sa morale, sa curiosité et sa réflexion. A partir de résultats qui peuvent sembler anodins, il est capable d'en décupler leur portée et de les rendre attrayants. Malgré un planning surchargé et face à mes contraintes diverses, il a consacré énormément de temps et d'énergie pour m'aider sur mes articles, de sorte que je puisse soutenir ma thèse dans les délais impartis. Au-delà de son aide à la rédaction, il a répondu à un grand nombre de mes interrogations sur la psychoacoustique et il m'a aidé à mieux comprendre mon sujet qu'est l'EAS. J'espère pouvoir prolonger les collaborations avec lui pour de longues années.

Merci à toute l'équipe d'Audition Conseil Lyon, à Patricia Elbaze, Renaud Gayte, Paul Berger, Jean Rouquet pour leur accueil, leur soutien et leur bonne humeur !

Merci au laboratoire NSCC de Gerland (et à Rémi Gervais), grâce à qui j'ai pu rencontrer les grosses pointures du domaine de l'audition et de l'implant cochléaire telles que Brian Moore, Robert Shannon, Norbert Dillier ou encore Colette McKay. Un grand Merci aussi aux secrétaires Manuella Allegro et Florence Léger.

Spéciale dédicace à Mikael Ménard, Idrick Akhoun et Jeanne Guiraud, les anciens du labo C. Je regrette vraiment de ne plus bosser avec eux. Du jour où ils m'ont accueilli, j'ai senti qu'ils deviendraient des super potes et effectivement nous avons vraiment partagé des moments inoubliables. Pour moi ce sont eux qui ont instauré l'ambiance et l'esprit chaleureux qui règne au Labo C.

J'adresse un grand merci aux autres étudiants qui sont passés par le Labo C : Kostas Kosta, Houda Saoud, Jean-Laurent Astier, Ludovic Bellier, Arnaud Jeanvoine, Céline Richard, Fathi Kallel, Julien Hoarau, Kamel Adjout, Julien Hanss, Cécile Barnel, Isabelle Ploin. C'était vraiment cool de travailler avec toutes ces personnes. Un merci tout particulier à Houssam Ibrahim, qui travaillait comme moi sur l'EAS, ça a été un réel plaisir de travailler avec lui et je lui souhaite de réussir dans sa carrière professionnelle car il le mérite.

Merci à tous les stagiaires GBM, ainsi qu'à Megha Agrawal, étudiante à l'Université de Vellore. Ce fut une expérience enrichissante de l'accueillir et de travailler avec elle.

Merci aux autres membres de l'équipe audition pour leur soutien : Annie Moulin, Michel Mazzuca, Rafael Laboissière. Et un grand merci à Evelyne Veuillet pour son aide et le temps qu'elle a consacré à mon dossier pour le comité d'éthique.

Je tiens à remercier tout le personnel du pavillon U, les services d'ORL, d'audiologie et d'explorations orofaciales. Parmi eux, l'équipe du réglage d'implants cochléaires, les audioprothésistes, orthophonistes et médecins Tiphaine Bigeard, Carine Bory, Muriel Kreiss, Veronique Ravoux, Jean-François Vesson, Solveig Chapuis, Brigitte Flamens, Adeline Croze, Anne-Marie Jonas, Corinne Fresco, Ines Aresi, Geneviève Lina-Granade, Hélène Apruzzese, Monique Porot, David Lamure, Pauline Garnier. Mais également les audiométristes et infirmières, Corinne Bruchet, Delphine Meunier, Sylvie Paquet, Pascale Melki, Nadine Monin, Laurence Cenni, et l'interne Vincent Pitiot. Et pour finir, les secrétaires Anick Fontbonne, Corinne Monceau, Florence Dambelin, Muriel Letondot, Sarah Renaud, Isabelle Perroud, Magalie, Marie-Pascale, Marie. Merci à tous pour votre aide, votre soutien, vos encouragements, pour la confiance que vous m'avez accordée et pour les super moments que j'ai partagés avec vous. Je n'oublierai jamais mon passage au pavillon U ni le repas surprise que vous m'avez préparé pour mes 30 ans ! J'essayerai de repasser de temps en temps, histoire qu'on me rappelle qu'il ne faut plus manger au McDo...

Un grand Merci à tous mes potes d'audioprothèse, sans doute les meilleurs amis que je me sois faits au cours de mes études post-bac, et sur qui j'ai pu compter pour venir passer mes tests à l'hôpital à toute heure du jour et de la nuit. Je pense en particulier à Marine Coustet, Julien Manuel, Nicolas Deville, Bérangère Guerrero, Yoann Juenet, Marine Gevaudan, Claudio Monteiro, Thomas Lelard, qui m'ont toujours fait oublier le stress du quotidien avec des fous rires interminables et des jeux aussi « débiles » que distrayants ! Je pense également à Christophe Bourg, Arnaud Hanus, Mathias Benoit, Valentine Moreau, Carole Becher, Philippe Perrin, Thomas Labaou, Florent Cotton, Denise Vogel, Madely Roux, Edouard Allermoz, Christophe Olivier, Sophie Negrello, Jean-Philippe Delbort, Guillaume Brun, Eddy Bouziane. Merci aussi à mes amis de Villefontaine : Benoit Laurent, Guillaume Illien, Gildas Guillaume, et les Janodi. A tous, je vous remercie pour votre soutien et pour toutes les soirées qu'on a passées ensemble. J'espère qu'on restera en contact encore de longues années.

Merci à Claire Chanal qui aura été bien plus qu'une binôme de stage.

Merci aux voisines Laurence Hert et Marie Grosdemouge qui m'ont toujours accueillies à bras ouverts dans leur appart'. Merci pour les petits plats que vous me prépariez quand je n'avais pas le temps de manger et pour toutes les petites attentions que vous aviez pour moi. Ca a fait un gros vide quand vous êtes parties du 23 rue Viala ! Je garde des super souvenirs de vous.

Merci également aux patients malentendants et implantés cochléaires qui ont accepté de participer à mes recherches.

Enfin, je voudrais terminer par les personnes les plus importantes, c'est-à-dire par mes parents et ma famille, qui ont toujours été présents, m'ont encouragé et m'ont aidé activement. Ils se sont sacrifiés et se sont pliés en quatre pour moi. Je leur en suis profondément reconnaissant et je ne les remercierai jamais assez pour ça ! J'espère que cette thèse sera une fierté pour eux.

SPÉCIFICITÉS DE L'IMPLANT ÉLECTRO-ACOUSTIQUE : INDICATIONS, INTERFACE BIOÉLECTRIQUE ET STRATÉGIE DE CODAGE

Le clinicien se trouve parfois confronté à des sujets qui présentent une surdité supérieure à 90 dB HL au-delà de 1 kHz avec une audition résiduelle dans les fréquences graves. Pour réhabiliter les hautes fréquences, il existe aujourd'hui différentes technologies : amplification conventionnelle, compression fréquentielle, implant cochléaire et depuis une dizaine d'année la stimulation électro-acoustique EAS qui consiste à stimuler acoustiquement les sons graves et électriquement les sons aigus via un implant cochléaire.

La première partie de cette thèse a consisté à identifier les facteurs qui influencent les capacités des patients sourds partiels à traiter l'information basse fréquence de la parole. Nous avons utilisé un test d'audiométrie vocale filtrée passe-bas. Nos résultats indiquent que les scores d'intelligibilité de la parole sont positivement corrélés avec la durée de la surdité. Ceci signifie qu'avec le temps, ces sujets malentendants apprennent à comprendre avec cette audition type filtre passe-bas, à tel point que certains ont des performances supra-normales pour l'utilisation des basses fréquences. Nos résultats montrent également une corrélation négative entre l'âge d'apparition de la surdité et les scores d'intelligibilité. Ce test pourra aider le clinicien à mieux cibler l'appareillage le plus adapté à chaque profil de patient.

La seconde partie de cette thèse, consacrée à l'EAS, a consisté à évaluer par des simulations chez le normo-entendant, diverses stratégies de codage du son par l'implant EAS. Actuellement, la stratégie utilisée pour l'EAS est calquée sur celle de l'implant cochléaire et nos résultats suggèrent que cette stratégie peut être optimisée.

SPECIFICITIES OF ELECTRIC-ACOUSTIC STIMULATION : INDICATIONS, BIOELECTRICAL INTERFACE AND CODING STRATEGY

Clinicians may face patients who have a deafness superior to 90 dB HL above 1 kHz with good low-frequency residual hearing. Today, several technologies are available to provide high frequencies: conventional amplification, frequency compression, cochlear implant since about 10 years Electric-Acoustic Stimulation EAS which consists in stimulating acoustically low frequencies while stimulating electrically high frequency sounds via a cochlear implant.

The first part of this dissertation consisted in identifying the factors which may influence abilities of partially deaf subjects to process low-frequency speech information. We used a low-pass filtered speech test. Our results show that speech intelligibility scores are positively correlated to the duration of deafness. This means that these hearing-impaired subjects learn to understand with this lowpass-like hearing, in such a way that some of them exhibit supranormal abilities for the processing of low-frequency sounds. Our results also show a negative correlation between the age at onset of deafness and speech intelligibility scores. This test may help the clinician to better evaluate which device would be best for every patient's profile.

The second part of this dissertation, about EAS, consisted in evaluating through simulations in normal-hearing listeners, several coding strategies by the EAS implant. Now the strategy used for EAS duplicates the strategy used by cochlear implants and our results suggest that this strategy could be optimized.

DISCIPLINE : Ingénierie pour le vivant

MOTS-CLES : Implantation cochléaire ; Stimulation Electrique-Acoustique ; Surdité partielle ; Parole filtrée passe-bas ; Intelligibilité de la parole ; Durée de privation ; Age d'apparition de la surdité ; Combinaison des informations ; Stratégie de codage ; Vocoder.

INTITULE ET ADRESSE DE L'U.F.R. OU DU LABORATOIRE :

INSERM U1028 - CNRS UMR5292 - Centre de Recherche en Neurosciences de Lyon

Centre Hospitalier le Vinatier, Batiment 452

95 Bd Pinel, 69675 Bron Cedex, 69500, France

Sommaire

I. Introduction : Surdit� et correction auditive	1
I.1. Rappels sur la parole	2
I.2. Rappels sur l'audition et la surdit�	3
I.3. Index d'articulation, audibilit�, mod�les pr�dictifs et concordance entre audiom�trie tonale et audiom�trie vocale	12
I.4. Les syst�mes de r�habilitation de la fonction auditive	13
I.4.1. L'aide auditive	14
I.4.2. L'implant cochl�aire	15
I.5. Evaluation des sourds r�habilit�s : De l'aide auditive � l'implant cochl�aire	17
I.6. Implant cochl�aire et intelligibilit� : Comment expliquer la disparit� des performances ?	18
I.7. Cas particulier des sujets qui ont des restes sur les graves	19
I.8. Appareillage des surdit�s dans les hautes fr�quences par les techniques de r�duction fr�quentielle	27
I.9. La stimulation Electrique-Acoustique (EAS)	28
I.10. Restitution de l'audibilit� des fr�quences sup�rieures � 1000 Hz dans le cas de surdit�s partielles	31
I.11. Variabilit� des performances en EAS	32
I.12. Bilan sur les performances avec les diff�rents syst�mes	34
I.13. Probl�matiques �mergentes pour l'EAS	35
I.14. Travaux r�alis�s et objectifs	37
 II. Productions scientifiques	 41
II.1. Mise au point d'un test de parole filtr�e pour mieux �valuer les indications de l'implantation EAS (article publi�)	42
II.2. Effet de l'�ge d'apparition de la surdit� et de la dur�e de privation sur les performances de compr�hension de la parole, dans le cas de surdit� dans les hautes fr�quences (article publi�)	44
II.3. Reconnaissance de la parole filtr�e passe-bas chez les sujets ayant une surdit� dans les hautes fr�quences (article soumis)	47

II.4. Suivi longitudinal de 8 patients EAS : expérience de Lyon	49
II.5. Simulation d'un implant EAS avec un vocoder hybride : Influence du nombre de canaux sur l'intelligibilité de la parole (article publié)	67
II.6. Analyse de « l'avantage de la stimulation combinée » par des modèles de combinaison de l'information (article sous presse)	69
II.7. Evaluation de différentes stratégies de codage de la parole par des simulations d'implant EAS (article soumis)	70
III. Discussion générale	72
III.1. Résumé des objectifs de la thèse	73
III.2. Résumé des principaux résultats de la thèse	74
III.3. Perspectives	75
IV. Références bibliographiques	84

I. Introduction :

Surdité et correction auditive

I.1. Rappels sur la parole

La parole est un stimulus de la fonction auditive. Elle peut se définir de façon simple, comme une succession de sons complexes périodiques (en général les voyelles = sons stationnaires) et/ou apériodiques (en général les consonnes = sons transitoires) qui, articulés les uns aux autres, vont avoir un sens et un aspect sémantique. Elle peut être décomposée en une série d'indices acoustiques, temporels et spectraux qui la caractérisent.

D'après Rosen (1992), l'information de parole peut être décomposée en 3 caractéristiques temporelles principales, qui véhiculent différents indices segmentaux :

- * L'enveloppe correspond aux fluctuations lentes de l'amplitude du signal, à des taux entre 2 et 50 Hz. L'enveloppe peut être principalement décrite par des caractéristiques acoustiques telles que l'intensité, la durée, l'attaque ou le temps de montée et le temps de retour.

- * La périodicité correspond aux propriétés du signal de parole qui relatent de la distinction entre stimulation périodique et apériodique. Les sons périodiques fluctuent généralement à des fréquences entre 50 et 500 Hz, tandis que les sons apériodiques fluctuent typiquement à des fréquences de quelques kHz jusqu'à 5–10 kHz.

- * La structure temporelle fine correspond aux fluctuations rapides d'une onde, à l'intérieur même de certaines périodes pour un son périodique, ou sur un intervalle de temps bref pour des sons apériodiques. Elle correspond en quelque sorte aux fréquences de modulations rapides du signal, dominant entre 600 Hz et 10 kHz.

La figure 1 issue de Loizou (2006), illustre une décomposition de la voyelle /a/ (à gauche) selon les indices définis par Rosen (1992). Le signal en haut à droite représente l'enveloppe du /a/ et le signal en bas à droite représente sa structure fine. Quant à l'information de périodicité elle se devine intuitivement, sur cette figure nous pouvons voir 3 périodes du /a/.

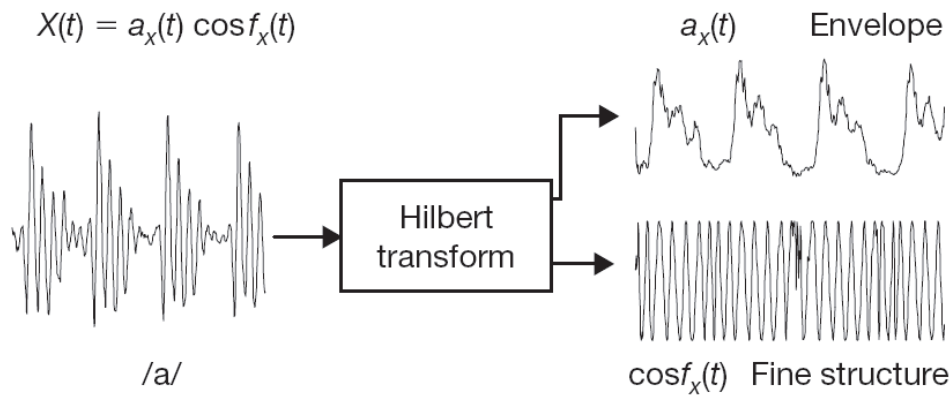


Figure 1. Décomposition d'un signal (ici une voyelle /a/) en enveloppe et structure fine par la transformée de Hilbert. D'après Loizou 2006.

L'enveloppe temporelle va donc représenter les modulations d'amplitude du signal sonore tandis que la structure temporelle fine va représenter les modulations de fréquence.

I.2. Rappels sur l'audition et la surdité

L'audition est une fonction sensorielle qui nécessite la réception, la transmission et l'intégration des messages sonores. La perception des sons se fait grâce à un ensemble d'éléments et de relais qui véhiculent l'information depuis les organes sensoriels périphériques (les oreilles) jusqu'au système auditif central (cortex auditif, cerveau).

La plupart du temps, la surdité est la conséquence de lésions situées au niveau périphérique. Ces lésions peuvent avoir des répercussions au niveau central qui vont se manifester par des modifications plastiques et/ou fonctionnelles. C'est ce que nous décrivons en partie dans la suite de ce manuscrit.

L'importance de la surdité s'évalue à l'aide de deux principaux examens cliniques simples. La perception des sons est évaluée par l'audiométrie tonale et la perception des messages sonores (= la parole) est évaluée par l'audiométrie vocale.

La perte auditive s'exprime en dB HL et la perte tonale moyenne (PTM) se calcule d'après les seuils de perception à 500, 1000, 2000 et 4000 Hz.

$$\text{Perte Tonale Moyenne} = (S_{500} + S_{1000} + S_{2000} + S_{4000}) / 4$$

La déficience auditive peut être classifiée en 6 grandes catégories (définies par le *Bureau International d'Audiophonologie*) :

- Audition normale : PTM = [0 – 20 dB HL]
- Déficience auditive légère : PTM = [21 – 40 dB HL]
- Déficience auditive moyenne : PTM = [41 – 70 dB HL]
- Déficience auditive sévère : PTM = [71 – 90 dB HL]
- Déficience auditive profonde : PTM = [91 – 119 dB HL]
- Déficience auditive totale : PTM = [120 dB HL]

Toutefois, la surdité ne se résume pas à de simples distorsions d'intensités et bien souvent la perte auditive est associée à des déficits plus complexes.

Quelques notions sur le déficit auditif et les principaux mécanismes de la surdité

Pour les pertes légères à moyennes, ce sont généralement les cellules ciliées externes (CCE) de la cochlée qui sont altérées. Au-delà de 60 dB HL de perte, pour une surdité neurosensorielle, les CCE sont généralement complètement détruites, mais également une partie des cellules ciliées internes (CCI) (Liberman et Dodds, 1984) (figure 2). Cela signifie que le système auditif ne peut plus bénéficier des mécanismes cochléaires actifs pour affiner la sélectivité fréquentielle. La conséquence directe est un élargissement des filtres auditifs, qui peut se manifester soit de part et d'autre de la fréquence centrale du filtre, soit principalement vers les basses fréquences. Lorsque les CCI sont endommagées, il en résulte une élévation des seuils et un élargissement du filtre de part et d'autre des cellules lésées. Dès lors on peut s'attendre, en stimulant l'audition avec un son pur, à percevoir non pas une fréquence mais une bande de fréquences autour de la fréquence caractéristique de ce son pur.

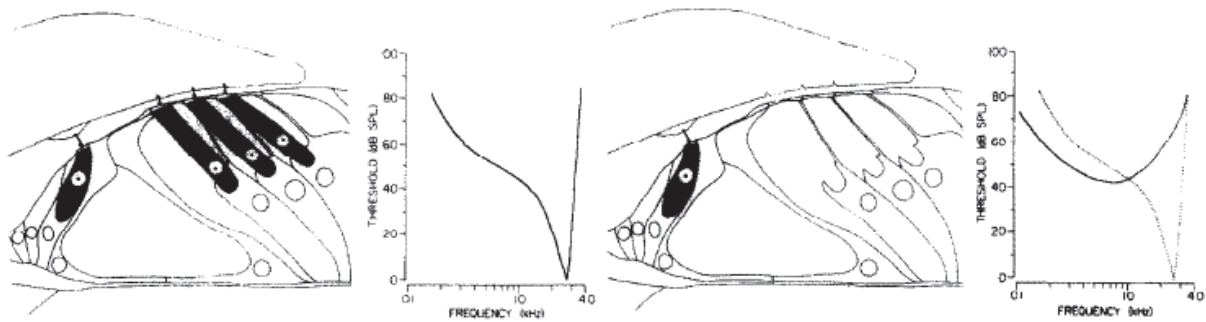


Figure 2. Représentation schématique d'un organe de Corti normal (à gauche) et d'un organe de Corti avec des CCE détruites (à droite). La partie droite montre la sélectivité fréquentielle (sous forme d'une courbe d'accord à la fréquence testée) pour chaque situation : lorsqu'il n'y a plus de CCE, les mécanismes cochléaires actifs n'agissent plus. D'après Liberman et Dodds, 1984.

Réduction de la dynamique auditive

La perte de sensibilité auditive correspond à une réduction de la dynamique auditive. C'est à dire que la surdité va se manifester par une élévation des seuils de perception mais également par une diminution des seuils d'inconfort. Ce phénomène s'appelle le recrutement de sonie, il se manifeste par un accroissement anormalement rapide de la sonie avec l'intensité de stimulation. L'origine physiologique du recrutement n'est pas vraiment connue, mais Yates (1990) a cependant proposé une explication intéressante à ce phénomène. Le rôle des CCE est d'amplifier la vibration de la membrane basilaire seulement pour les niveaux de stimulation faible. De ce fait, la fonction entrée-sortie de la membrane basilaire dans une cochlée saine est non-linéaire pour les faibles niveaux d'entrée, mais linéaire pour les plus fortes intensités. Plus simplement, l'amplification diminue lorsque l'intensité d'entrée augmente. Dans une cochlée privée de la fonction des CCE, le déplacement de la membrane basilaire est une fonction linéaire du niveau d'entrée, quel que soit le niveau (figure 3). Si on considère que la sonie est proportionnelle aux déplacements de la membrane basilaire, alors aux faibles niveaux d'entrée, une cochlée sans l'amplification des CCE aura besoin d'un niveau de stimulation plus forte pour obtenir un déplacement de la membrane basilaire de la même ampleur que celui d'une cochlée qui aurait un système de CCE intègre. Les niveaux maximums de confort restent quant à eux identiques voire diminuent, d'où la réduction de la dynamique observée.

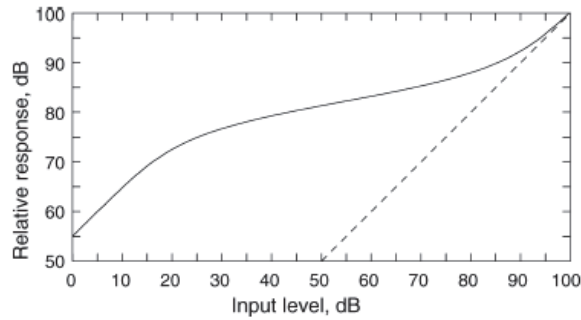


Figure 3. Fonction schématique d'entrée-sortie de la membrane basilaire pour une sinusoïde à la fréquence caractéristique. Une échelle décibel est utilisée pour les 2 axes. L'ordonnée est échelonnée (arbitrairement) de telle sorte qu'une entrée de 100 dB procure une sortie de 100 dB. La ligne en pointillés montre la pente qui serait obtenue si la réponse était linéaire. D'après Moore, 2007.

Elargissement des filtres auditifs

La cochlée agit comme un analyseur de spectre. Par analogie aux filtres électroniques, les filtres auditifs sont caractérisés par leur bande passante qui peut être définie comme la bande fréquentielle à l'intérieure de laquelle la réponse est maximale et en dehors de laquelle la réponse est inférieure d'au moins 3 dB par rapport à la réponse maximale. Si nous prenons par exemple un filtre dont la réponse pic est centrée à 1000 Hz et dont la réponse est atténuée de 3 dB pour les fréquences inférieures à 900 Hz et supérieures à 1120 Hz, alors la bande passante de ce filtre sera 220 Hz.

La bande passante d'un filtre auditif peut être déterminée par ce qu'on appelle la bande passante rectangulaire équivalente (ERB) qui est décrite par l'équation :

$$ERB_N = 24.7(0.00437F + 1)$$

L'indice N étant utilisé pour préciser que cette valeur d'ERB (Equivalent Rectangular Bandwidth) est valable pour des normo-entendants, F étant la fréquence centrale du filtre. F et l'ERB sont exprimées en Hz (Glasberg et Moore, 1990).

La sélectivité fréquentielle correspond à la capacité à discriminer deux sons éloignés en fréquence. Pour le codage tonotopique, si deux sons appartiennent au même filtre auditif (s'ils sont fréquentiellement proches) alors l'oreille ne sera pas capable de les séparer. Si au contraire les deux sons sont éloignés en fréquence, de sorte qu'ils appartiennent à deux filtres auditifs différents, alors l'oreille les percevra comme étant deux sons distincts. Ce phénomène est valable pour le codage tonotopique uniquement, mais il existe d'autres facteurs temporels (phénomène de battement, possible tant qu'il y a du verrouillage de phase ou phase locking) qui permettent de discriminer 2 sons. Il est à noter que tous les filtres auditifs de la cochlée n'ont pas la même bande passante. Comme le suggère l'équation qui définit l'ERB, plus la fréquence centrale du filtre augmente, plus la largeur du filtre augmente (figure 4, partie gauche). Naturellement, la sélectivité fréquentielle est donc moins bonne dans les hautes fréquences que dans les basses fréquences.

La sélectivité fréquentielle est également liée à l'intensité de stimulation. Plus l'intensité est élevée, plus la membrane basilaire sera soulevée et plus la jupe des filtres auditifs sera large. La figure 4 (partie droite) illustre la forme d'un filtre auditif centré à 1 kHz pour différentes intensités de stimulation.

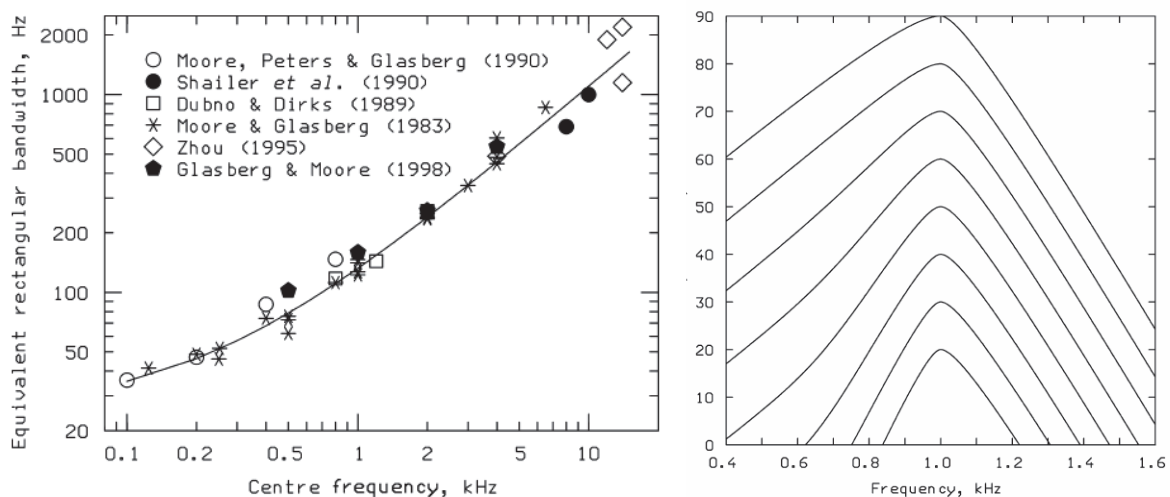


Figure 4. Partie gauche : estimation de la bande passante des filtres auditifs, en fonction de la fréquence centrale, d'après une série d'expérimentations. Partie droite : forme du filtre auditif centré à 1 kHz, tracé pour des intensités d'entrée de 20 à 90 dB SPL/ERB_N par pas de 10 dB. D'après Moore, 2007.

Chez le malentendant, les lésions des CCE entraînent un élargissement des filtres cochléaires, ce qui implique une diminution de la résolution fréquentielle (Zwicker et Schorn, 1978 ; Florentine et al., 1980 ; Glasberg et Moore, 1986). Les filtres auditifs sont d'autant plus larges que la perte auditive est importante. Selon Moore (2007), les largeurs des filtres auditifs deviennent 2 fois et 4 fois supérieure à la normale, respectivement pour des pertes auditives moyennes (41 – 70 dB) et sévères (71 – 90 dB).

La figure 5, extraite de Moore (2007), montre le pattern d'excitation généré par une voyelle « i », pour une oreille normale et pour différentes conditions de filtres auditifs d'oreilles lésées. On peut supposer que si la voyelle est présentée en isolation dans le silence, elle peut être identifiable par le malentendant dans la mesure où l'information sur les formants est partiellement préservée. Mais la différence d'amplitudes relatives entre les pics et les creux spectraux (= le détail spectral) se dégrade avec la perte auditive et l'élargissement des filtres auditifs. De ce fait, les voyelles deviennent probablement de moins en moins identifiées lorsque la perte auditive devient de plus en plus importante.

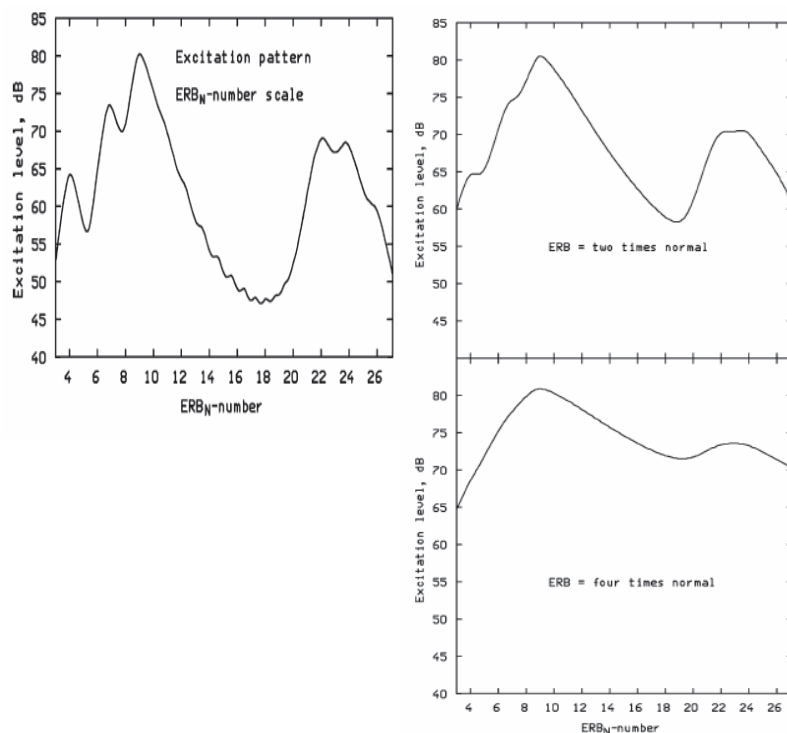


Figure 5. Pattern d'excitation généré par une voyelle « i », pour une oreille normo-entendante (à gauche), pour une oreille avec des filtres auditifs 2 fois plus larges que la normale (en haut à droite), et pour une oreille avec des filtres auditifs 4 fois plus larges que la normale (en bas à droite). D'après Moore, 2007.

Pour évaluer, l'effet de l'élargissement des filtres auditifs chez le malentendant, des études ont proposé des simulations chez le normo-entendant (Baer et Moore, 1993). Pour simuler la perte de sélectivité fréquentielle, des signaux de parole, présentés dans le silence et en situation de bruit concurrent ont subi un lissage spectral. Les auteurs ont mesuré la reconnaissance de ces signaux dégradés par des sujets normo-entendants, pour différentes conditions de lissage (pas de lissage vs élargissement de l'ERB d'un facteur 3 vs élargissement de l'ERB d'un facteur 6) et de bruit (calme vs RSB = 0 dB vs RSB = -3 dB). Leurs résultats montrent que pour la reconnaissance de phrases dans le calme, l'intelligibilité n'est que légèrement affectée par le lissage et la parole reste très intelligible (>90%) malgré des filtres 6 fois plus larges que la normale. En revanche, en situation de bruit concurrent (speech-shaped noise), l'intelligibilité des phrases est très affectée par le lissage, en particulier pour les lissages très importants et pour les RSB très bas, dès que le bruit devient supérieur à la parole (RSB = -3 dB) (figure 6).

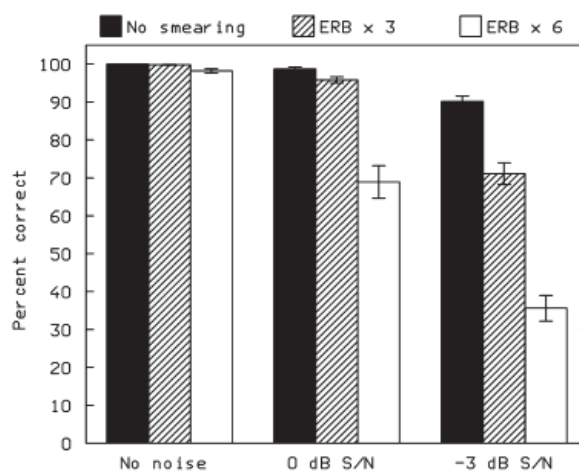


Figure 6. Pourcentage d'intelligibilité des phrases, pour différentes conditions de lissage spectrale et différents niveaux de bruit. D'après Baer et Moore (1993).

Ces données suggèrent qu'en milieu calme la parole reste relativement intelligible malgré une information fréquentielle très appauvrie (Baer et Moore, 1993). D'où l'idée que les malentendants se baseraient essentiellement sur des indices temporels (tels que l'enveloppe du signal) pour comprendre. Les déficits liés à la perte auditive auraient donc un impact essentiellement sur le traitement fréquentiel, mais n'auraient qu'un impact modéré sur le traitement temporel. Il a d'ailleurs été montré que les malentendants auraient des capacités de traitement temporel normales ou supérieures à la normale (Fullgrabe et al., 2003). De même, pour la perception de la hauteur tonale, le malentendant utilise plus (ou essentiellement) les indices d'enveloppe temporelle par rapport à un normo-entendant (Oxenham, 2008). L'hypothèse suivante proposée par Heinz et al. (2010) pourrait expliquer ces meilleures capacités de traitement temporel par les malentendants. Dans une cochlée pathologique, les filtres auditifs sont élargis, ce qui a pour conséquence le recrutement d'un plus grand nombre de fibres nerveuses à l'intérieur de ces filtres. L'élargissement des filtres auditifs va produire une augmentation de la corrélation entre les réponses d'une série de fibres auditives adjacentes à l'intérieur de ce filtre, de part et d'autre de la fréquence caractéristique du filtre. Il y aura alors beaucoup plus de fibres synchronisées aux fluctuations d'amplitude du stimulus, ce qui expliquerait les capacités de traitement temporel normales ou meilleures que la normale.

Conséquences de la perte auditive sur le traitement des informations d'enveloppe et de structure fine

Comme nous l'avons défini précédemment, la structure temporelle fine va représenter les modulations de fréquence. Dans une cochlée pathologique qui a perdu ses propriétés de sélectivité fréquentielle, on peut supposer que l'information de structure fine ne sera pas traitée. Si les filtres auditifs sont très larges, et que l'oreille a perdu sa capacité à discriminer des sons séparés en fréquence, alors elle ne sera probablement pas sensible aux modulations de fréquence véhiculées par l'information de structure fine (figure 7). Là encore, le fait de ne plus avoir accès à l'information de structure fine va avoir pour conséquence d'importants déficits pour l'écoute dans le bruit.

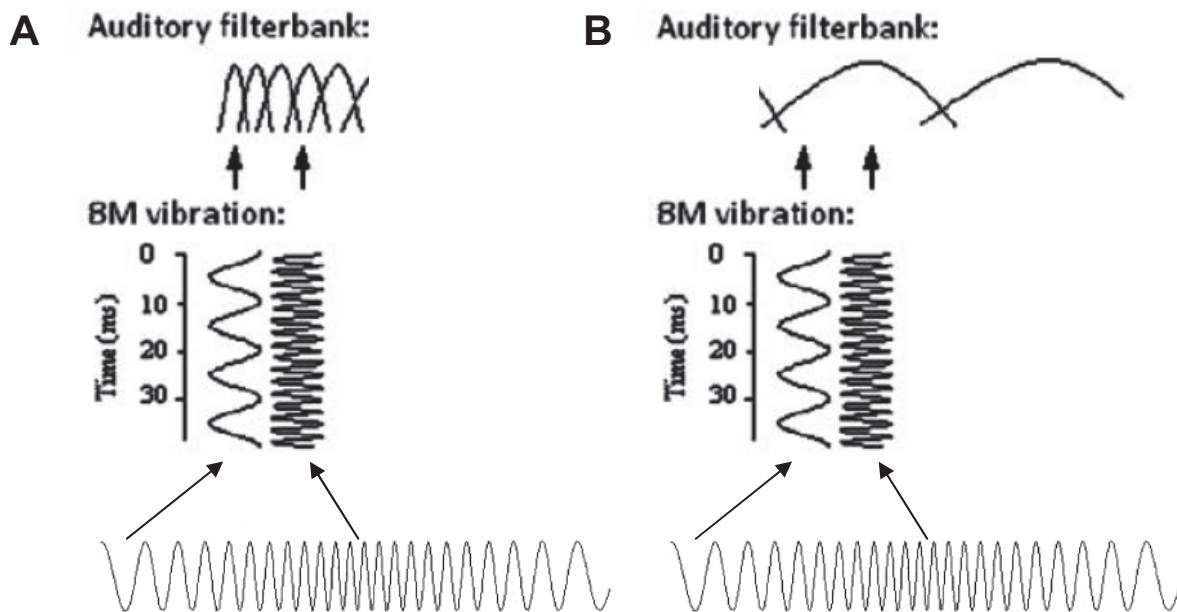


Figure 7. En bas : signal modulé en fréquence représentant un exemple de structure fine d'un signal. Au milieu : réponse de la membrane basilaire aux instants où le signal est de basse fréquence et de plus haute fréquence. En haut : filtres auditifs correspondants à ces fréquences caractéristiques dans une cochlée saine (A) et dans une cochlée lésée qui aurait ses filtres auditifs élargis (B). Dans la situation (B), le sujet ne sera pas capable de discriminer les variations de fréquence.

Pour la perception de la parole dans un bruit de fond fluctuant, les normo-entendants tirent un bénéfice considérable de l'écoute dans les vallées temporelles et spectrales du bruit, c'est-à-dire là où le rapport signal bruit (cible-masqueur) local est élevé. Cette capacité à écouter dans les vallées résulterait de la perception de la structure fine (Hopkins et al., 2008). A la différence des normo-entendants, des études psychophysiques récentes suggèrent que les malentendants auraient des capacités réduites pour l'utilisation des indices de structure fine (figure 8). Leur mauvaise compréhension dans les bruits de fond fluctuants (type cocktail party) serait liée à l'incapacité à traiter la structure fine (Lorenzi et al., 2006 ; Hopkins et Moore, 2007 ; Hopkins et al., 2008).

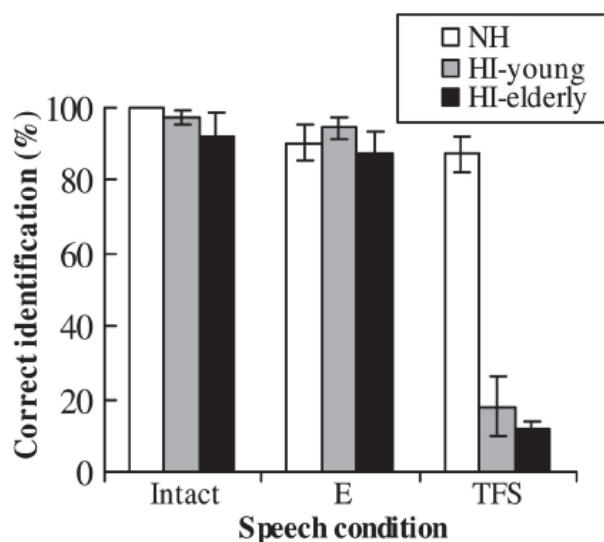


Figure 8. Pourcentage de reconnaissance des consonnes par 3 groupes de sujets (NH = normo-entendants, HI-young = malentendants jeunes, HI-elderly = malentendants plus âgés). Les consonnes étaient présentées soit en condition normale (Intact), soit en condition où seule l'information d'enveloppe est préservée (E), soit en condition où seule l'information de structure fine est préservée (TFS). D'après Lorenzi et al. (2006).

Lorsque la perte auditive devient trop importante, il peut arriver qu'un son (généralement fort) qui stimule cette zone lésée soit audible alors que la région est morte. En réalité, le son est détecté à un endroit où la vibration de la membrane basilaire sera moins importante, mais où les CCI et les neurones fonctionnent plus efficacement. Moore et al. en 2000, ont mis au point un test qui permet de détecter de telles régions mortes ou inertes de la cochlée : le TEN test (Threshold Equalizing Noise = bruit égaliseur de seuil). Le terme de « Zones mortes Cochléaires » (ZIC) a été introduit en 2000 par Moore, pour qualifier une région de la cochlée dans laquelle les CCI et/ou les neurones fonctionnent tellement peu/mal qu'un son pur produisant un pic de vibration dans cette région est détecté par des cellules répondant habituellement à une autre fréquence caractéristique. D'après Moore (2001), une zone où la perte est supérieure à 90 dB HL dans les hautes fréquences, peut-être associée à une zone morte cochléaire.

I.3. Index d'Articulation, audibilité, modèles prédictifs et concordance entre audiométrie tonale et audiométrie vocale

Index d'Articulation, audibilité, modèles prédictifs et concordance entre audiométrie tonale et audiométrie vocale

(Article soumis aux Cahiers de l'audition)

Fabien Seldran (1,2), Doctorant en Ingénierie Biomédicale, audioprothésiste.

Lionel Collet (1), Professeur des universités-Praticien Hospitalier, Chef du service d'audiologie et d'explorations oro-faciales des Hospices Civils de Lyon, Directeur de l'école d'Audioprothèse de Lyon.

Stéphane Gallégo (1,3), Docteur en Ingénierie Biomédicale, audioprothésiste.

(1) Hôpital Edouard Herriot, Pavillon U, service d'audiologie et d'explorations oro-faciales, Lyon

(2) Vibant Med-El, Sophia-Antipolis

(3) Audition Conseil, Lyon

L'audiométrie se décline en 2 épreuves principales et complémentaires : l'audiométrie tonale et l'audiométrie vocale. L'audiométrie tonale est un examen qui consiste à évaluer le seuil d'audition d'un sujet pour différentes fréquences, classiquement entre 250 et 8000 Hz qui couvrent l'essentiel des sons de la parole. Toutefois l'audiométrie tonale, aussi précise soit-elle, fournit une information quantitative sur le degré de la surdité, mais elle ne permet pas d'évaluer l'impact de la surdité sur la compréhension du patient dans sa vie quotidienne. Il est nécessaire pour cela de réaliser un test complémentaire : l'audiométrie vocale, qui indique le pourcentage de reconnaissance de la parole à différentes intensités de stimulation.

Il existe plusieurs matériels phonétiques qui diffèrent les uns des autres par leur composition et leur structure, chacun permettant de mettre en évidence des troubles de l'audition plus ou moins spécifiques. Dans le domaine de l'audiologie nous pouvons trouver des listes de logatomes, des listes de nombres, des listes de mots ou des listes de phrases. En fonction du matériel, l'unité phonétique utilisée pour évaluer le score (pourcentage de reconnaissance) peut être le phonème, le mot, le nombre, les mots clés dans une phrase, le phonème cible (par exemple la consonne présentée dans un contexte VCV). Ces matériels phonétiques font plus ou moins appel à la suppléance mentale du sujet. Les tests peuvent être réalisés dans le silence ou avec un bruit concurrent type bruit blanc, bruit rose (bruit filtré), bruit de parole (speech

noise = bruit qui présente les mêmes caractéristiques spectrales que la parole), ou encore type cocktail party.

Du fait de la multitude des matériels phonétiques, nous pouvons constater des différences de performances parfois chez un même sujet, avec 2 tests différents et parfois avec 2 listes d'un même test. Ces variations peuvent s'expliquer par des différences entre les listes utilisées qui peuvent être dues par exemple à des différences de difficulté ou encore à des différences de composition spectrale. Ceci est vrai pour le normo-entendant mais aussi et surtout pour le malentendant.

Pour évaluer fidèlement le déficit d'un patient malentendant, le clinicien doit être capable d'apprécier si le patient est dans la normale, ou s'il est sous-performant ou sur-performant, par rapport aux sons qu'il perçoit et par rapport au matériel phonétique utilisé. Il est donc important de pouvoir juger la capacité du patient à utiliser les informations qui lui sont fournies pour comprendre.

Il existe pour cela des modèles mathématiques élaborés et développés depuis les années 50's, qui permettent de prédire l'intelligibilité d'un sujet en fonction de l'information qu'il perçoit. Ces modèles ont été utilisés dans de nombreux travaux pour évaluer dans un premier temps quels étaient les sons utiles à la compréhension de la parole et pour quantifier la proportion d'information fournie par différentes bandes de fréquences. Dans un deuxième temps, les modèles ont été réutilisés et adaptés à plusieurs reprises pour mesurer l'intelligibilité en fonction de l'audibilité chez les malentendants.

Caractéristiques spectrales de la parole

Pour pouvoir prédire l'intelligibilité, il est important de connaître les caractéristiques de la parole qui vont être utiles à la compréhension, tant au niveau des informations d'intensités que des informations spectrales.

Pour cela, Byrne et al. (1994) ont étudié le spectre moyen de la parole à long terme dans 12 langues. Ils ont enregistré des locuteurs de différentes nationalités qui avaient pour consigne de lire un passage simple d'un livre à haute voix, dans un microphone situé à 20 cm d'eux. L'intensité moyenne qu'ils ont mesurée était de 71,8 dB SPL pour les voix masculines et de

71,5 dB SPL pour les voix féminines, ce qui correspond à une intensité de 58 dB pour la voix normale à 1 mètre. La dynamique de la parole quant à elle varie de 30 à 50 dB selon les études (Studebaker et al., 1999 ; Zeng et al., 2002).

En ce qui concerne le spectre de la parole, il ressort des mesures de Byrne et al. (figure 1) que la composition fréquentielle est sensiblement la même pour les différentes langues et que l'essentiel de l'énergie du signal est généralement concentrée sur une plage de fréquences de 250 à 700 Hz environ avec une intensité à peu près constante. A 1kHz l'énergie chute d'environ 10 dB par rapport à l'intensité des basses fréquences et à 2kHz la diminution est de 15 dB.

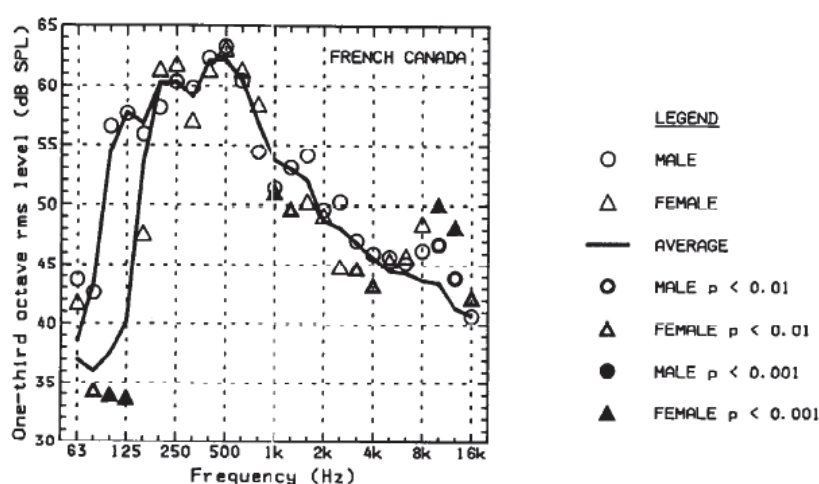


Figure 1 : Spectre moyen à long terme de la parole, pour la langue française (Français Canadien). D'après Byrne et al. (1994).

D'après ces mesures, nous pourrions penser que les sons médiums et aigus sont modérément importants pour la compréhension, dans la mesure où ils sont beaucoup moins présents dans le spectre à long terme de la parole. En raisonnant ainsi, on peut supposer qu'un malentendant, privé de la perception des hautes fréquences (au-delà de 1 ou 2 kHz) n'est que moyennement gêné pour comprendre, puisqu'il perçoit les sons basse fréquence qui véhiculent une grande partie (voire l'essentiel) de l'énergie. Mais tout n'est pas aussi simple, en effet l'information véhiculée par une bande de fréquence n'est pas proportionnelle à l'énergie contenue dans cette bande de fréquence. En réalité certaines bandes de fréquences sont plus importantes que d'autres pour la perception de la parole. Ceci a été mis en évidence

avec des outils mathématiques tels que les fonctions d'importance fréquentielle et l'index d'articulation.

Fonctions d'importance fréquentielle

En 1987, Studebaker et al. ont étudié les fonctions d'importance fréquentielle pour du discours continu quotidien. Selon eux, la bande la plus importante pour la compréhension de la parole se situerait entre 400 et 500 Hz et la bande autour de 2500 Hz serait relativement importante aussi. D'autre part ils comparent leurs résultats avec les fonctions d'importance fréquentielle reportées dans 3 autres études qui utilisent des autres matériels phonétiques (figure 2) et qui trouvent que les bandes les plus importantes seraient situées dans des fréquences plus hautes. D'après la fonction de l'ANSI S3.5-1969, la bande la plus importante serait centrée vers 2 kHz. Tandis que pour French et Steinberg (1947) et Black (1959), elle se situerait à 2500 Hz. Turner et al. (1998) ont également étudié les fonctions de pondération fréquentielle sur 4 bandes de fréquences et leurs résultats montrent de façon reproductible, que la bande la plus importante serait la bande [0-1120] Hz suivie de la bande [2250-3500] Hz, ce qui est à peu près cohérent avec ce qui a été observé dans les études précédentes. En France, les travaux de Gilbert et Michey (2005) ont mis en évidence que la gamme de fréquences [100-750] Hz était la moins affectée par le bruit (cocktail party) et relativement aux autres bandes, elle devient la bande la plus importante en milieu bruyant par rapport à un milieu calme.

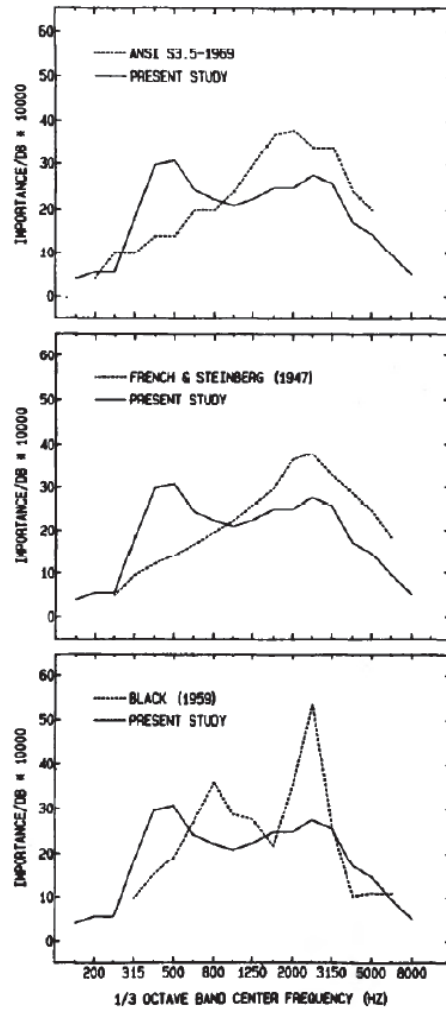


Figure 2. Fonction d'importance fréquentielle (importance / dB) mesurée au 1/3 d'octave, pour du discours continu mesuré par Studebaker et al. (1987), comparées aux fonctions d'importance de l'ANSI S3.5-1969, et aux valeurs dérivées des études de French et Steinberg (1947) avec des syllabes sans signification et pour Black (1959) avec des mots monosyllabiques. D'après Studebaker et al., 1987.

La connaissance de ces fonctions d'importance fréquentielle s'avère alors très utile pour les fabricants d'appareils auditifs, en particulier pour leurs systèmes de débruitage. En effet, les appareils auditifs peuvent calculer le RSB dans chacun de leurs canaux et le principe de ces systèmes de débruitage est de réduire l'amplification des bandes fréquentielles qui contiennent du bruit. En connaissant l'importance des différentes bandes, le système pourra pondérer la réduction de gain maximale autorisée dans chaque bande. La réduction du gain dans les régions fréquentielles qui ont moins d'importance affecte moins l'intelligibilité qu'une réduction de gain dans les régions avec une plus grande importance (figure 3). C'est à

dire qu'à atténuation égale du gain, on perdra moins d'informations pour les bandes qui contribuent peu à la compréhension, par rapport à des bandes qui contribuent beaucoup à la compréhension.

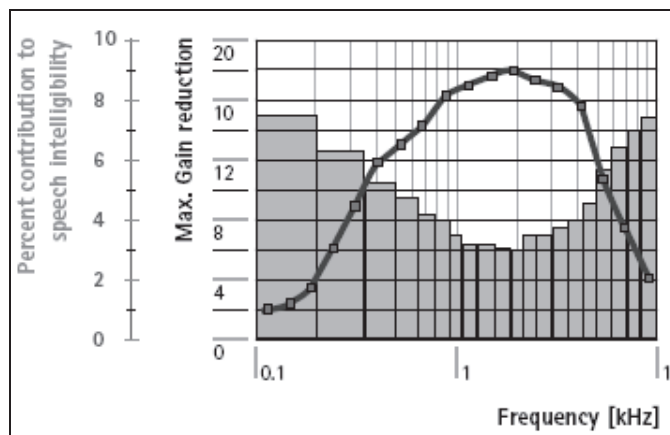


Figure 3. Atténuation maximale du gain, autorisée par un système de débruitage, en fonction de la contribution des différentes bandes fréquentielles pour l'intelligibilité de la parole. Adapté de Pavlovic, 1984 (source Phonak).

Index d'Articulation

Dans le domaine de l'audiologie, il est important de pouvoir vérifier la concordance entre l'audiométrie tonale et l'audiométrie vocale en ayant, dans la mesure du possible, des mesures précises et sensibles. Pour cela, les acousticiens ont développé un outil appelé l'index d'articulation, c'est une fonction qui a été élaborée dans le but de prédire la reconnaissance vocale d'un sujet en fonction des sons perçus ou transmis, i.e. de l'audibilité.

Initialement, l'index d'articulation et ses variantes ont été créés pour déterminer l'intelligibilité des sons de parole transmis par le téléphone, ce n'est qu'un peu plus tard qu'elles ont trouvé une application en audiologie. L'intelligibilité est prédite à partir des différentes caractéristiques de la parole, telles que le spectre à long terme, les modulations d'amplitude ou encore l'importance des différentes bandes fréquentielles pour la compréhension de la parole (mesuré par masquage ou filtrage de certaines bandes de fréquences).

Les méthodes les plus appliquées pour la prédiction de l'intelligibilité de la parole sont l'Index d'Articulation (AI = Articulation Index) (French et Steinberg, 1947 ; Fletcher et Galt,

1950 ; ANSI, 1969), qui a par la suite été modifié en Index d'Intelligibilité de la Parole (SII = Speech Intelligibility Index) (ANSI, 1997), puis en Index de Transmission de la Parole (STI = Speech Transmission Index) (Houtgast and Steeneken, 1985). La puissance de ces modèles vient de la quantité de connaissances empiriques sur lesquelles ils reposent. Tous ces modèles supposent que la parole est codée par plusieurs canaux fréquentiels qui véhiculent une information indépendante. Toutefois l'indépendance des bandes n'est qu'un postulat arbitraire pour le calcul de l'index car en réalité, la plupart du temps il se produit des interactions de synergie et de redondance entre les différents canaux.

L'index d'articulation peut être gouverné par l'équation :

$$AI = \sum_i AI_i,$$

Avec AI correspondant à l'index d'articulation cumulatif de tous les canaux et AI_i correspondant à l'index d'articulation d'un canal isolé.

L'AI et le SII sont dérivés du signal de parole en calculant le rapport signal/bruit (RSB ou SNR) dans les différentes bandes fréquentielles :

$$AI = \sum_i \frac{W_i(SNR_i + 15)}{30},$$

W_i étant un facteur de pondération du canal fréquentiel i et SNR_i correspondant au rapport signal/bruit dans ce même canal i . W_i dépend du matériel utilisé et prend en compte le fait que les hautes fréquences sont plus importantes pour la reconnaissance des consonnes que pour la reconnaissance d'une phrase. Les principales différences entre les différentes versions d'AI et SII sont la manière dont ils prennent en compte les non-linéarités telles que les distorsions, le masquage et l'élargissement des bandes fréquentielles.

Contrairement à l'AI qui est déterminé à partir du spectre du matériel phonétique et du RSB, le STI utilise la « modulation transfer function » et il est particulièrement efficace pour prédire l'intelligibilité de la parole dans les auditoriums et dans les pièces, car il prend en compte le lissage de l'information véhiculée par l'enveloppe de parole (c'est-à-dire la

réduction de la profondeur de modulation). Le STI a l'avantage d'être applicable pour tous les types de bruits et de distorsions qui affectent le signal de parole, que ce soit les réverbérations ou les autres aberrations qui peuvent se produire au cours du temps.

Le passage de l'AI, du SII, ou du STI, à l'intelligibilité de la parole demande une transformation non-linéaire qui doit être adaptée aux données empiriques. La transformation va dépendre du type de matériel phonétique utilisé et la fonction est généralement plus raide pour le matériel avec contexte (par exemple des phrases) que pour le matériel avec des mots isolés (figure 4). En clair, à quantité égale d'information transmise, un auditeur comprendra plus facilement une phrase qu'un mot sans signification.

D'après Hill et al. (1968), ce que nous pouvons également observer sur la figure 4, si un auditeur est capable de discriminer 70% de phonèmes ou de mots monosyllabiques isolés, alors il est capable de comprendre une phrase, une conversation, grâce au contexte et à la suppléance mentale.

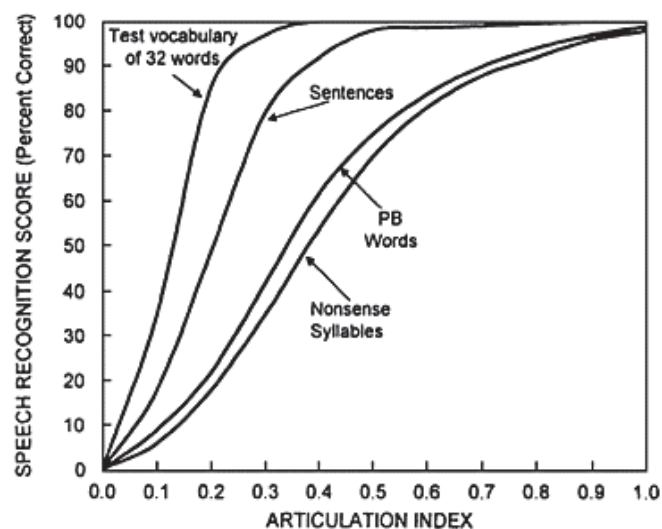


Figure 4. Relation entre l'index d'articulation et la compréhension de la parole pour différents matériels phonétiques. (Source Gelfand, 2005)

Comme nous l'avons précisé un peu plus tôt, les différentes bandes fréquentielles ne sont pas totalement indépendantes. Ce qui se manifeste parfois par une sous-prédiction ou une sur-prédiction des performances par les index, lorsque plusieurs bandes spectrales sont combinées. En effet la présentation simultanée de deux bandes spectrales ne se résume pas en une simple additivité des informations véhiculées par chacune des bandes. En fonction de leur

proximité, les bandes vont fournir des informations plus ou moins redondantes, et plus ou moins complémentaires, voire synergiques.

Modèles d'intégration de l'information provenant de plusieurs sources

Le succès des modèles basés sur l'index d'articulation laissent supposer qu'il est possible d'utiliser les mesures de performances dans des conditions de bandes isolées pour prédire les scores dans les conditions de bandes (ou de modalités) combinées pour des sujets individuels, sans avoir nécessairement besoin d'utiliser des paramètres tels que les fonctions d'importance fréquentielle. Plusieurs de ces modèles sont décrits notamment dans les travaux de Braidà (1991) et de Ronan et al. (2004). Il existe des modèles simples de combinaison des bandes, c'est par exemple le cas du modèle de sommation de probabilité. Mais il existe également des modèles plus complexes, c'est par exemple le cas des modèles de perception à logique floue (FLMP = Fuzzy Logical Model of Perception, Massaro, 1987) ou encore des modèles d'intégration pré- ou post-étiquetage (Braidà, 1991). Bien entendu, ces modèles supposent également qu'il y a des incertitudes et des « bruits » qui peuvent se manifester à différents étages du traitement. Ces bruits peuvent être externes (par exemple bruit de fond, locuteur concurrent) ou internes (par exemple fluctuation de l'attention du sujet). Ils peuvent être indépendants (i.e. ils se produisent avant l'étape d'intégration) ou tardifs (i.e. ils se produisent après l'étape d'intégration). Un modèle donné va être plus ou moins optimal dans ses prédictions en fonction des paramètres qu'il prend en compte et des postulats sur lesquels il est fondé.

Selon Macmillan (1987), la réalisation d'une tâche d'identification ou de discrimination peut être ramenée à un processus décisionnel. Pour Massaro (1987), la perception nécessite différentes étapes : une étape sensorielle dans laquelle le sujet va évaluer les « traits », une étape d'intégration des traits et une étape décisionnelle.

Dans un premier temps le sujet va d'abord évaluer les différentes sources d'information de manière indépendante. Chacune des informations (valeurs, indices caractéristiques) va être convertie en un ensemble de propriétés qui s'appellent des « traits ». Chacun de ces traits va donner une valeur qui est plus ou moins proche et plus ou moins ressemblant à des prototypes connus par le sujet (par exemple à son lexique, aux phonèmes, aux syllabes ou aux mots de son vocabulaire). La seconde étape, l'intégration, consiste à combiner (à intégrer) les traits

issus des différentes sources et à faire correspondre cette information combinée à des prototypes connus par le sujet. Enfin l'étape de décision consiste à attribuer comme réponse le prototype qui présente le maximum de ressemblance avec l'information sensorielle résultant de l'étape d'intégration.

La figure 5, issue de Braida (1991), représente des espaces d'indices hypothétiques pour des expériences d'identification des plausives /p, b, k, g/ en condition de stimulation audiovisuelle. L'axe horizontal symbolise une source d'information visuelle, la double flèche verticale délimite cet espace en 2 parties : consonnes bilabiales /p/ et /b/ à gauche, versus consonnes vélares /k/ et /g/ à droite. L'axe vertical symbolise une source d'information auditive, la double flèche horizontale délimite cet espace en 2 parties : consonnes sourdes /p/ et /k/ en bas, versus consonnes voisées /g/ et /b/ en haut. Pour cette tâche d'identification, un sujet soumis à la stimulation visuelle seule pourra facilement catégoriser si la consonne est bilabiale ou vélaire mais il aura une très grande incertitude pour catégoriser si cette consonne est voisée ou non. Réciproquement, la stimulation auditive seule permettra de catégoriser si la consonne est voisée ou non avec une incertitude pour catégoriser si cette consonne est bilabiale ou vélaire. Après avoir délimité ces sous espaces, nous obtenons les 4 prototypes qui serviront de points de comparaisons après l'étape d'intégration des informations auditives et visuelles.

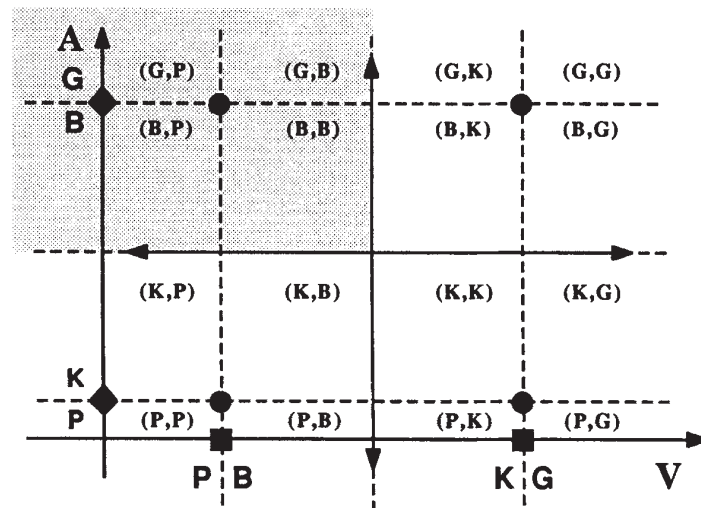


Figure 5. Espaces d'indices hypothétiques pour une tâche d'identification /p, b, k, g/. Axe horizontal : indices scalaires pour la modalité visuelle. Axe vertical : indices scalaires pour la modalité auditive. D'après Braida, 1991.

L'exemple présenté ci-avant illustre un modèle d'intégration de 2 source : une source d'information visuelle et une source d'information auditive. Ce même modèle peut être utilisé de façon analogue pour prédire l'intelligibilité lorsque 2 sources auditives (par exemple 2 bandes fréquentielles spectralement distantes) sont présentées simultanément. C'est ce que nous décrivons dans le paragraphe suivant.

Complémentarité et redondance des informations

En 1995, Warren et al. ont évalué sur 420 sujets, l'intelligibilité de phrases (pourcentage de mots-clés identifiés) filtrées passe-bande à des fréquences centrales de 370, 530, 750, 1100, 1500, 2100, 3000, 4200 et 6000 Hz avec des largeurs de bandes de 1/3 d'octave (expérience 1) ou de 1/20 d'octave (expérience 2) et des pentes abruptes (96 dB/octave pour les bandes larges d'1/3 d'octave et 115 dB/octave pour les bandes de largeur 1/20 d'octave).

Leurs résultats montrent qu'hormis pour les 2 bandes extrêmes (370 et 6000 Hz), l'intelligibilité reste relativement bonne, voire très bonne (figure 6). Pour les deux bandes extrêmes, leur intelligibilité est très faible lorsqu'elles sont diffusées individuellement, voire quasi-nulle dans certains cas. En revanche lorsque ces deux bandes sont diffusées simultanément, elles ont un effet super-additif et le score en condition dual-bande est significativement supérieur à la somme des scores en condition bande individuelle (figure 7).

Pour l'expérience 1, en condition 1/3 d'octave, l'intelligibilité fournie par les bandes centrées à 370 Hz et à 6000 Hz est respectivement de 23% et 24%. Lorsque ces deux bandes sont combinées, l'intelligibilité devient 78% en condition diotique et 76% en condition dichotique.

De même, pour l'expérience 2, en condition 1/20 d'octave, l'intelligibilité fournie par les bandes centrées à 370 Hz et à 6000 Hz est respectivement de 0,9% et 10%. Et lorsque les deux bandes sont combinées en condition diotique, l'intelligibilité moyenne atteint 27,8%.

Cela signifie que l'information contenue dans des bandes distantes peut être intégrée pour produire une augmentation de l'intelligibilité qui est beaucoup plus importante que la simple additivité.

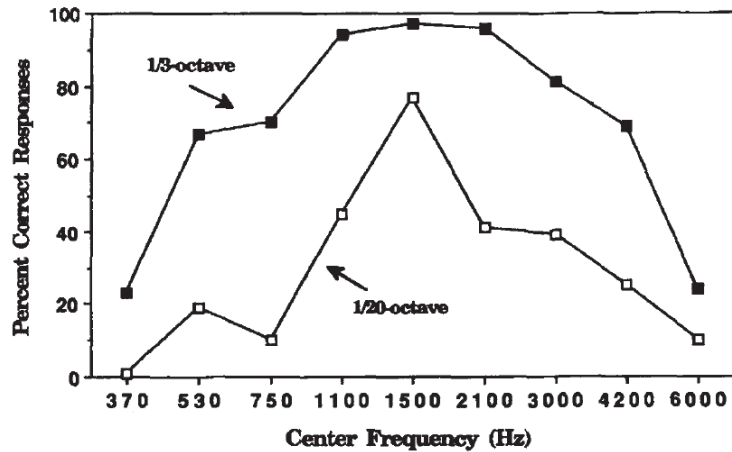


Figure 6. Pourcentage moyen de mots clés identifiés, pour des phrases (CID sentences) présentées en condition largeur de bande 1/3 d'octave et en condition largeur de bande 1/20 d'octave, pour neuf fréquences centrales. D'après Warren et al., 1995.

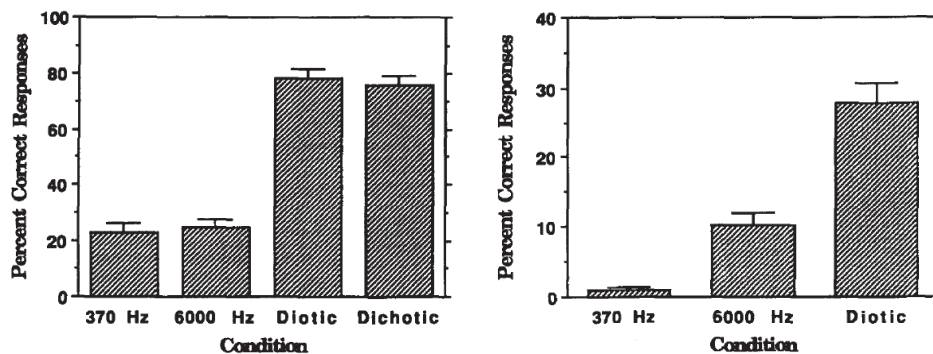


Figure 7. Pourcentage moyen de mots clés identifiés pour des phrases filtrées passe-bande (1/3 d'octave à gauche, 1/20 d'octave à droite, de fréquences centrales 370 et 6000 Hz), présentées soit individuellement, soit simultanément. D'après Warren et al., 1995.

En 1996, Lippmann a observé un effet similaire. Il propose dans son étude de faire écouter des consonnes qui ont subi un filtrage passe-bas à 800 Hz auquel il ajoute ou non l'information d'un passe haut à 3,15, 4, 5, 6,3, 8, et 10 kHz, afin de laisser un trou spectral dans les stimuli.

L'information fournie en condition passe-bas seul fournit une intelligibilité moyenne de 44,3%. Lorsque cette information est couplée avec de l'information haute-fréquence, l'intelligibilité atteint 52,7% si le passe-haut commence à 10 kHz, 73,9% pour une fréquence de coupure à 8 kHz et jusqu'à 91,6% avec une fréquence de coupure à 3,15 kHz.

Ces observations renforcent l'idée qu'il existe des indices redondants dans les différentes régions fréquentielles. Une forte dégradation du signal d'un point de vue quantitatif, ne résulte pas forcément en une dégradation équivalente sur le plan qualitatif : pour un signal privé de l'information 800 Hz–4000 Hz, Lippmann observe une intelligibilité proche de 90%. L'hypothèse d'une synergie ou d'une super-additivité de l'information provenant de bandes spectralement éloignées serait donc confirmée.

Grant et Braida (1991) avaient conduit, quelques années auparavant, des expériences analogues. Leur test consistait à faire écouter à un groupe de sujets normo-entendants, des phrases phonétiquement équilibrées, assez difficiles, dans lesquelles ils devaient reconnaître 5 mots clés, en condition audio seul et audio-visuel. Dans une des sessions de test, les sujets écoutaient les stimuli filtrés passe-bande de largeur 1/3 d'octave. Les bandes étaient présentées individuellement ou combinées 2 à 2 : Bande 1 [80–630 Hz] ; Bande 2 [800–1000 Hz] ; Bande 4 [2500 Hz] ; Bande 5 [3150 Hz] ; Bande 1 + Bande 2 ; Bande 4 + Bande 5 ; Bande 1 + Bande 4 ; Bande 1 + Bande 5 ; Bande 2 + Bande 5.

En condition audio seul (figure 8), leurs résultats expérimentaux montrent que les performances sont largement sous-estimées par la théorie de l'Index lorsqu'ils combinent deux bandes fréquentielles distantes (-41%), et inversement leurs résultats sont surestimés lorsque les deux bandes fréquentielles sont adjacentes (+18%).

Pour les résultats sur-prédits en condition de bandes adjacentes, ils émettent l'hypothèse qu'il se produit un phénomène de masquage de l'information entre les bandes. Et ce phénomène de masquage serait d'autant plus important que les bandes sont proches l'une de l'autre. L'autre explication qu'ils proposent, c'est qu'il existe probablement un plus grand degré de corrélation entre des bandes adjacentes qu'entre des bandes éloignées. Lorsqu'on ajoute des bandes 1/3 d'octave successives, les « nouveaux » indices fournis seraient en fait redondants avec l'information contenue dans les bandes plus basses. Ce qui aurait pour conséquence de fournir seulement une légère augmentation de l'intelligibilité. Au contraire, le fait de combiner des bandes 1/3 d'octave non-adjacentes fournirait des indices complémentaires, ce qui résulterait en une amélioration importante de l'intelligibilité.

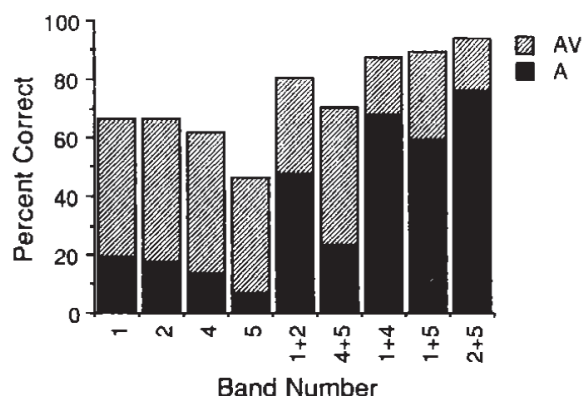


Figure 8. Pourcentage de mots clés reconnus en fonction des bandes spectrales présentées, en conditions audio seule (A) et audio-visuelle (AV). D'après Grant et Braida, 1991.

Application des modèles prédictifs chez le malentendant

D'après l'index d'articulation et ses variantes, l'intelligibilité doit augmenter lorsque l'audibilité augmente. En réalité ceci est surtout vrai pour les sujets normo-entendants et pour les malentendants dont la perte auditive est légère ou moyenne, mais dès que la perte auditive dépasse un certain degré, les prédictions des modèles ne sont pas toujours parfaites. Evidemment, pour les populations malentendantes il est nécessaire de prendre en compte les seuils audiométriques des individus testés afin de calculer les valeurs AI_i de l'index. Si l'index prédisait fidèlement les performances et s'il ne nécessitait aucune modification, cela signifierait que l'audibilité est le seul facteur limitant l'intelligibilité. Or, la plupart des études ont montré que l'intelligibilité mesurée chez les malentendants était moins bonne que l'intelligibilité prédite par l'AI (Fletcher, 1952; Dugal et al., 1978; Pavlovic, 1984; Pavlovic et al., 1986; Smoorenburg, 1992), en particulier pour les sujets qui ont des pertes modérées ou sévères. Les données recueillies chez le normo-entendant sont souvent homogènes, ce qui permet d'obtenir des modèles assez fidèles. Lorsque nous sommes en présence d'oreilles pathologiques, les données sont plus dispersées et il existe de larges variations inter-individuelles, à audibilité équivalente (Magnusson, 1996, figure 9 partie gauche). Sherbecoe et Studebaker (2003) ont montré que du fait de cette variabilité, les modèles perdent de leur précision et il devient nécessaire, pour gagner en précision, que ces modèles soient adaptés et prennent en compte des facteurs correctifs (par exemple un facteur de distorsion due à l'intensité de la parole, le facteur âge, ou encore le facteur désensibilisation qui réfère à une réduction des performances qui ne peut pas être expliquée par une réduction de l'audibilité).

La figure 9 (partie droite : 2^{ème} et 3^{ème} cadres) illustre par exemple des données recueillies par Pavlovic (1984) pour 2 groupes de surdités. Leur matériel phonétique était des listes de mots présentées dans diverses conditions de filtrage (large-bande, passe-bas et passe-haut), avec ou sans bruit blanc masquant à un RSB de 10 dB. Seize sujets malentendants ont été testés. Huit avaient des pertes moyennes (seuils meilleurs que 50 dB HL à 4000 Hz), le score moyen de ces sujets pour les 8 différentes conditions, indiquées par les nombres dans le 2^{ème} cadre de la figure 9, sont proches des prédictions de l'AI (courbe en trait plein). Pour les 8 autres sujets ayant des pertes plus importantes (seuils de 55 dB HL ou au-delà, à 4000 Hz), les scores moyens, présentés dans le panel droit de la figure 9, tombent quasiment toujours en dessous des scores prédits par l'index.

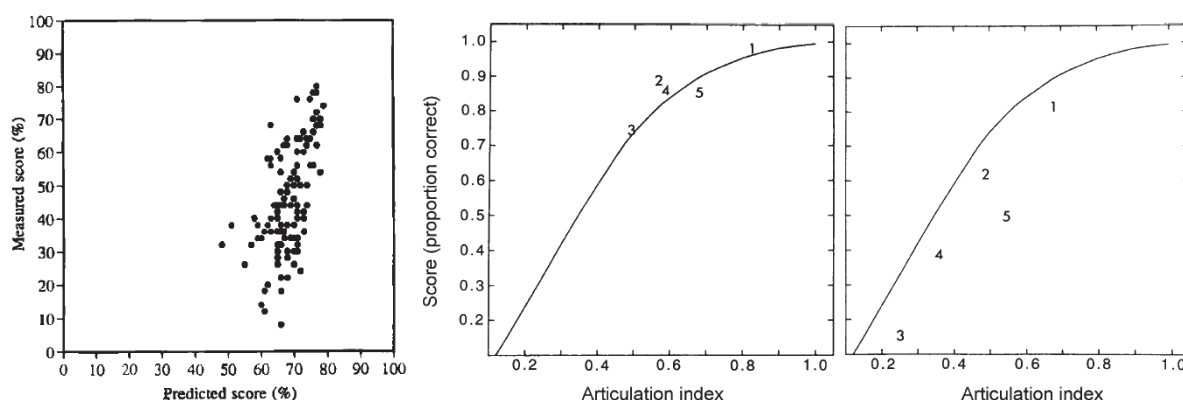


Figure 9 : Partie gauche : Relation entre les scores de reconnaissance de la parole prédits par le SII et mesurés chez 57 sujets malentendants (114 oreilles). D'après Magnusson, 1996. Partie droite (2^{ème} et 3^{ème} cadre) : Résultats de Pavlovic (1984) comparant les scores de reconnaissance de la parole de sujets malentendants avec les scores prédits sur la base de l'AI. Chaque nombre représente le score moyen des sujets pour une condition spécifique de filtrage/bruit de fond. Pour les sujets qui ont une perte modérée les prédictions sont précises (2^{ème} cadre) ; pour les sujets qui ont des pertes plus sévères, les scores mesurés sont inférieurs aux scores prédits (3^{ème} cadre). D'après Moore, 2007.

Dans le cas de surdité dans les hautes fréquences, une série de travaux réalisés à la fin des années 90's ont mis en évidence que le fait de rendre audible les hautes fréquences pouvait être contreproductif, en particulier pour les zones où la perte auditive dépasse 55-60 dB de perte (Hogan et Turner, 1998 ; Ching et al., 1998 ; Turner et Cummings, 1999). C'est-à-dire que l'amplification de ces fréquences ne va pas nécessairement améliorer la compréhension et dans certains cas elle peut la dégrader (Hogan et Turner, 1998 ; Vickers et al., 2001). Pour ces

patients, les modèles vont grandement surprédire les performances obtenues (Ching et al., 1998). Les patients sont donc sous-performants par rapport à ce que l'on pourrait espérer sur la base de l'audibilité. En général, les résultats des études qui ont utilisé l'AI ou le SII suggèrent que, si l'audibilité est un facteur essentiel, ce n'est pas le seul facteur impliqué dans la compréhension, au moins pour les malentendants qui présentent des pertes modérées à sévères. On peut se demander si ces modèles prédictifs sont applicables aux surdités partielles. Comment expliquer alors les différences entre les performances prédites et les performances observées ?

Les travaux de Egan et Hake (1950) et de Ehmer (1959a et 1959b) ont montré que les basses fréquences d'un son avaient un fort pouvoir masquant sur les composantes hautes fréquences. Cet effet de masquage est naturellement présent du fait de la composition du spectre moyen à long terme de la parole et il est accentué par les différences de seuils de perception entre les sons graves et les sons aigus, ainsi que par l'élargissement des filtres auditifs. Ce phénomène est appelé le masquage ascendant (*upward spread of masking*) et il s'explique assez simplement. L'onde provoquée par un son aigu va se propager essentiellement sur la partie basale de la cochlée et ne va pas engendrer de distorsion sur les basses fréquences. Tandis qu'une onde de basse fréquence qui va se propager tout au long de la cochlée, depuis la base jusqu'à l'apex, risque de provoquer des interférences au niveau basal. Pour cette raison, les composantes de basse fréquence d'un son vont dégrader la perception des composantes de haute fréquence, d'autant plus si les cellules ciliées externes (CCE) sont endommagées et si la stimulation est forte.

Pour une surdité neurosensorielle, si les seuils de perception sont meilleurs que 60 dB HL, ce sont essentiellement les CCE qui vont être endommagées. On suppose alors qu'une perte auditive de degré moyenne à sévère résulte en une perte de la fonction des CCE, avec une fonction des cellules ciliées internes (CCI) qui n'est pas nécessairement altérée. Les CCE sont plus fragiles et plus sensibles, par exemple aux agents oto-toxiques, que les CCI. On peut donc considérer qu'au-delà de 60 dB HL de perte les CCE sont complètement détruites et les CCI commencent à être endommagées à leur tour (Lieberman et Dodds, 1984). En absence de CCE, le sujet ne peut plus bénéficier des mécanismes cochléaires actifs, ce qui pourrait expliquer la valeur charnière à 60 dB observée dans les travaux de Hogan et Turner (1998), de Ching et al. (1998) et de Turner et Cummings (1999). La perte de la fonction des CCE a deux conséquences directes : des distorsions d'intensité qui se manifestent par une perte de

sensibilité et des distorsions de fréquence qui se manifestent par un élargissement des filtres auditifs.

Diverses études ont mesuré la forme des filtres auditifs chez des sujets qui avaient des lésions cochléaires. Les résultats montrent généralement que chez le malentendant, les filtres auditifs sont beaucoup plus larges que la normale et que la largeur augmente avec la perte auditive, dès que la perte dépasse 30 dB. Classiquement, pour une perte auditive moyenne (41 – 70 dB), la largeur des filtres auditifs (ERB) est 2 fois supérieure à la normale. Pour une perte sévère (71 – 90 dB), l'ERB est environ 4 fois supérieur à la normale (Moore, 2007). Les lésions au niveau des CCE entraîneraient ainsi un élargissement des filtres cochléaires, ce qui impliquerait la diminution de résolution fréquentielle (Zwicker et Schorn, 1978 ; Florentine et al., 1980 ; Glasberg et Moore, 1986). Le masquage et la diminution de résolution fréquentielle sont 2 exemples de facteurs qui pourraient expliquer les sur-prédictions des modèles, mais il existe probablement d'autres déficits perceptifs liés à la surdité.

Conséquences de la surdité sur les fonctions d'importance fréquentielle

Intéressons-nous à présent aux conséquences potentielles de la surdité sur les fonctions d'importance fréquentielle. En 2006, Hornsby et Ricketts ont évalué la contribution des basses fréquences pour la compréhension chez les sujets ayant une surdité neurosensorielle. Un groupe de 10 sujets malentendants et un groupe de 10 sujets contrôles normo-entendants ont été soumis à des tests de reconnaissance de parole filtrée passe-bas et passe-haut à différentes fréquences de coupure, filtrée large bande ou filtrée passe-bande.

Pour l'évaluation des performances, les auteurs ont utilisé la méthode de la fréquence d'intersection (crossover frequency). Cette méthode a été proposée par French et Steinberg en 1947. La fréquence d'intersection correspond à la fréquence de coupure pour laquelle un filtre passe-bas et un filtre passe-haut fournissent la même intelligibilité, c'est-à-dire que la fréquence d'intersection va séparer le spectre en deux bandes fréquentielles d'égale contribution pour la reconnaissance de la parole. Si cette fréquence d'intersection est basse par rapport à la normale, cela signifie que le sujet utilise plutôt ses basses fréquences pour comprendre la parole. Si au contraire elle est plus haute que la normale, cela signifie que le sujet utilise plus ses hautes fréquences pour comprendre. Par exemple, French et Steinberg ont déterminé que pour les normo-entendants cette fréquence d'intersection était de 1900 Hz pour

du matériel phonétique type syllabes CVC, mais cette fréquence n'est pas fixe, elle peut varier en fonction du matériel phonétique utilisé et probablement de la langue.

Les résultats obtenus par Hornsby et Ricketts sont présentés figure 10. Les performances des malentendants sont généralement inférieures à celles de normo-entendants, en particulier pour les conditions passe-haut. Leurs résultats montrent une différence significative de fréquence d'intersection entre les deux groupes ainsi que des différences de performance à ces fréquences. Le groupe normo-entendant a une fréquence d'intersection à 1408 Hz qui fournit une intelligibilité de 50%. La fréquence d'intersection du groupe malentendant est obtenue à 1207 Hz et fournit une intelligibilité de 28%. D'après les auteurs, cette différence de fréquence d'intersection serait principalement due au fait que les malentendants ont des performances réduites dans les conditions de filtrage passe-haut.

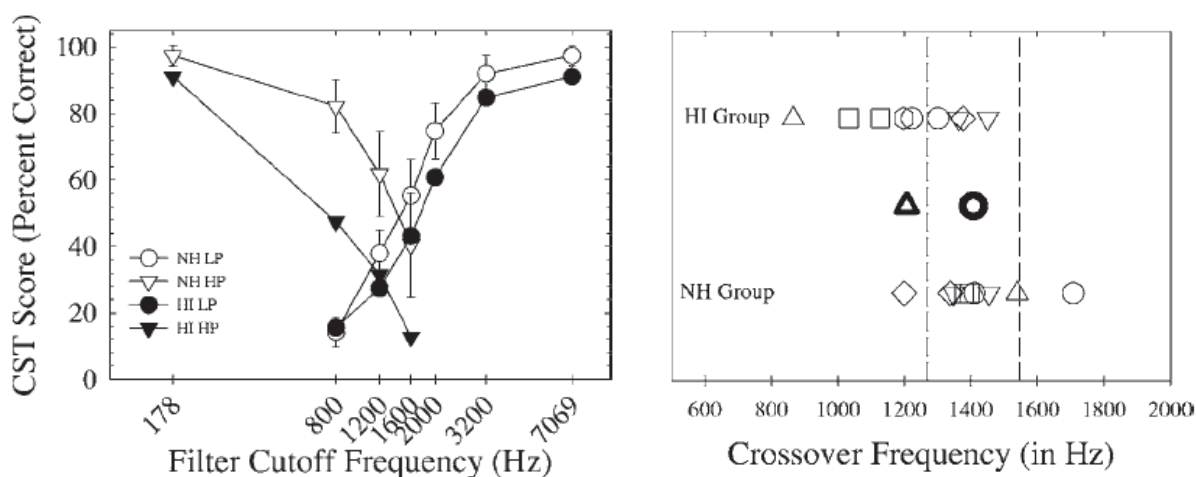


Figure 10. Partie gauche : Intelligibilité de la parole filtrée passe-bas et passe-haut à différentes fréquences de coupe, pour des normo-entendants et des malentendants ayant une perte en pente. Partie droite : fréquence d'intersection (données individuelles en traits fins et moyennes en traits gras) pour les deux groupes de sujets testés normo- vs malentendants. Les barres verticales représentent plus ou moins une déviation standard autour de la fréquence d'intersection des normo-entendants.

Pour confirmer cette hypothèse, les auteurs ont comparé les performances de leurs malentendants avec les prédictions calculées par le SII. Pour la condition passe-bas, les performances obtenues sont assez fidèles aux prédictions de l'index, pour les fréquences de coupe inférieures à 1200 Hz. Mais en condition de filtrage passe-haut, le SII surprédit les

SPÉCIFICITÉS DE L'IMPLANT ÉLECTRO-ACOUSTIQUE : INDICATIONS, INTERFACE BIOÉLECTRIQUE ET STRATÉGIE DE CODAGE

Le clinicien se trouve parfois confronté à des sujets qui présentent une surdité supérieure à 90 dB HL au-delà de 1 kHz avec une audition résiduelle dans les fréquences graves. Pour réhabiliter les hautes fréquences, il existe aujourd'hui différentes technologies : amplification conventionnelle, compression fréquentielle, implant cochléaire et depuis une dizaine d'année la stimulation électro-acoustique EAS qui consiste à stimuler acoustiquement les sons graves et électriquement les sons aigus via un implant cochléaire.

La première partie de cette thèse a consisté à identifier les facteurs qui influencent les capacités des patients sourds partiels à traiter l'information basse fréquence de la parole. Nous avons utilisé un test d'audiométrie vocale filtrée passe-bas. Nos résultats indiquent que les scores d'intelligibilité de la parole sont positivement corrélés avec la durée de la surdité. Ceci signifie qu'avec le temps, ces sujets malentendants apprennent à comprendre avec cette audition type filtre passe-bas, à tel point que certains ont des performances supra-normales pour l'utilisation des basses fréquences. Nos résultats montrent également une corrélation négative entre l'âge d'apparition de la surdité et les scores d'intelligibilité. Ce test pourra aider le clinicien à mieux cibler l'appareillage le plus adapté à chaque profil de patient.

La seconde partie de cette thèse, consacrée à l'EAS, a consisté à évaluer par des simulations chez le normo-entendant, diverses stratégies de codage du son par l'implant EAS. Actuellement, la stratégie utilisée pour l'EAS est calquée sur celle de l'implant cochléaire et nos résultats suggèrent que cette stratégie peut être optimisée.

SPECIFICITIES OF ELECTRIC-ACOUSTIC STIMULATION : INDICATIONS, BIOELECTRICAL INTERFACE AND CODING STRATEGY

Clinicians may face patients who have a deafness superior to 90 dB HL above 1 kHz with good low-frequency residual hearing. Today, several technologies are available to provide high frequencies: conventional amplification, frequency compression, cochlear implant since about 10 years Electric-Acoustic Stimulation EAS which consists in stimulating acoustically low frequencies while stimulating electrically high frequency sounds via a cochlear implant.

The first part of this dissertation consisted in identifying the factors which may influence abilities of partially deaf subjects to process low-frequency speech information. We used a low-pass filtered speech test. Our results show that speech intelligibility scores are positively correlated to the duration of deafness. This means that these hearing-impaired subjects learn to understand with this lowpass-like hearing, in such a way that some of them exhibit supranormal abilities for the processing of low-frequency sounds. Our results also show a negative correlation between the age at onset of deafness and speech intelligibility scores. This test may help the clinician to better evaluate which device would be best for every patient's profile.

The second part of this dissertation, about EAS, consisted in evaluating through simulations in normal-hearing listeners, several coding strategies by the EAS implant. Now the strategy used for EAS duplicates the strategy used by cochlear implants and our results suggest that this strategy could be optimized.

MOTS-CLES : Implantation cochléaire ; Stimulation Electrique-Acoustique ; Surdité partielle ; Parole filtrée passe-bas ; Intelligibilité de la parole ; Durée de privation ; Age d'apparition de la surdité ; Combinaison des informations ; Stratégie de codage ; Vocoder.

performances pour toutes les conditions de filtrage, ce qui suggère que ces malentendants auraient une capacité réduite à traiter l'information haute-fréquence.

Hornsby et Ricketts ont par ailleurs comparé les résultats de leur groupe malentendant avec ceux obtenus dans une précédente étude (Hornsby et Ricketts, 2003), chez un groupe de malentendants qui avaient une perte plate, uniforme sur toutes les fréquences. Les résultats de ces 2 groupes sont présentés sur la figure 11, elle représente la fonction de régression moyenne basée sur leurs performances mesurées en condition de filtrage passe-bas et passe-haut. Les fonctions de régression pour les données passe-haut se chevauchent pour les 2 groupes, excepté pour la condition large bande. Au contraire pour les données passe-bas, la fonction de régression montre des résultats plus mauvais pour les patients qui ont une perte plate, ce qui résulte en une fréquence d'intersection plus basse pour le groupe avec une surdité en pente, par rapport au groupe avec une perte plate. D'après les auteurs, la différence de fréquence d'intersection ne serait pas due à des différences de capacités à utiliser l'information haute fréquence. Selon eux, la fréquence d'intersection plus haute observée chez le groupe à perte plate reflèterait une moins bonne capacité à traiter l'information basse fréquence. On peut supposer que le groupe qui a une surdité en pente arrivera mieux à exploiter l'information basse fréquence par rapport à l'autre groupe.

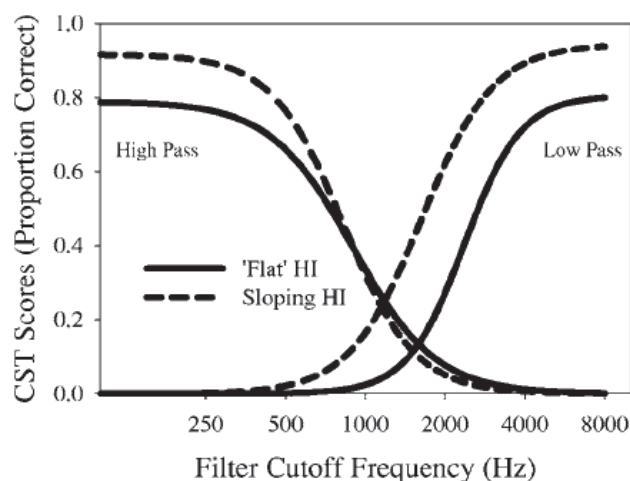


Figure 11. Fonctions de régressions moyennes basées sur les performances obtenues en conditions de filtrage passe-bas et passe-haut des groupes à perte plate (trait continu) et en pente (pointillés). D'après Hornsby et Ricketts (2006).

Dans un article de 2003, Vestergaard avait fait un constat similaire. L'auteur mesurait l'intelligibilité de sons de parole filtrés en passe-bas à différentes fréquences de coupure, chez des sujets présentant des zones inertes cochléaires (ZIC : Moore, 2001) dans les hautes fréquences. Il apparaît que les sujets dont l'audiogramme présente une fréquence de coupure basse sont capables de mieux utiliser l'information de parole basse fréquence par rapport à des sujets qui ont une valeur de la fréquence de coupure plus haute. Les patients dont la fréquence de coupure est basse auraient également, en condition d'audibilité faible, des meilleures capacités à traiter l'information basse fréquence par rapport à des sujets qui ne présentent pas de ZIC. Le même constat a été fait plus récemment par Moore et Vinay (2009).

Selon Hornsby et Ricketts (2006), les patients avec une perte en pente de ski sous-exploitent les hautes fréquences au profit des basses fréquences, d'où le décalage de la fréquence d'intersection vers les basses fréquences. Au contraire Horwitz et al. (2008) pensent que les malentendants ne tirent pas le même avantage des indices de parole audibles que les normo-entendants, toutefois ils seraient plus efficaces pour extraire l'information à laquelle ils ont été habitués, dans la zone fréquentielle où leur audition est la meilleure. Cette idée est également défendue par Vestergaard (2003).

Chez le malentendant, au-delà de la surdité et de l'audibilité, il y a probablement des mécanismes plus complexes qui entrent en jeu, fonctions de l'importance de la perte auditive et des caractéristiques du patient (par exemple la désensibilisation évoquée par Sherbecoe et Studebaker, 2003). Dans certains cas, pour gagner en précision et pour être applicables au malentendant, les modèles requièrent l'application de facteurs correctifs. Ces index, initialement élaborés pour le développement des télécommunications, s'avèrent très utiles dans le domaine de l'audiologie. Malheureusement ils sont encore peu connus des professionnels de l'audition et méritent d'être démocratisés.

Références :

- ANSI. ANSI S3.5-1969. American national standard methods for the calculation of the articulation index. American National Standards Institutes, New York 1969.
- ANSI. ANSI S3.5-1997. American national standard methods for the calculation of the speech intelligibility index. American National Standards Institutes, New York 1997.
- Black JW. Equally contributing frequency bands in intelligibility testing. *J Speech Hear Res.* 1959;2: 81-3.
- Braida LD. Crossmodal integration in the identification of consonant segments. *Q J Exp Psychol A.* 1991;43:647-77.
- Byrne D, Dillon H, Tran K, Arlinger S, Wilbraham K, Cox R, Hagerman B, Hetu R, Kei J, Lui C, Kiessling J, Kotby MN, Nasser NHA, El Kholy WAH, Nakanishi Y, Oyer H, Powell R, Stephens D, Meredith R, Sirimanna T, TavartkiladzeG, Frolenkov GI, Westerman S, Ludvigsen C. An international comparison of long-term average speech spectra. *J Acoust Soc Am* 1994;96:2108-20.
- Ching TY, Dillon H, Byrne D. Speech recognition of hearing-impaired listeners: predictions from audibility and the limited role of high-frequency amplification. *J Acoust Soc Am.* 1998;103:1128-40.
- Dugal R, Braida LD, Durlach NI. Implications of Previous Research for the Selection of Frequency-Gain Characteristics, in *Acoustical Factors Affecting Hearing Aid Performance* (eds Studebaker and Hochberg), University Park Press, Baltimore, 1978:379-403.
- Egan JP, Hake HW. On the masking pattern of a simple auditory stimulus. *J Acoust Soc Am.* 1950;22: 622-30.
- Ehmer RH. Masking patterns of tones. *J Acoust Soc Am.* 1959a;31:1115-20.
- Ehmer RH. Masking by tones vs. Noise bands. *J Acoust Soc Am.* 1959b;31:1253-6.
- Fletcher H. The perception of sounds by deafened persons. *J Acoust Soc Am.* 1952;24:490-7.
- Fletcher H, Galt RH. The perception of speech and its relation to telephony. *J Acoust Soc Am.* 1950;22:89-151.
- Florentine M, Buus S, Scharf B, Zwicker E. Frequency selectivity in normally-hearing and hearing-impaired observers. *J Speech Hear Res.* 1980;23:646-69.
- French NR, Steinberg JC. Factors governing the intelligibility of speech sounds. *J Acoust Soc Am.* 1947;19:90-119.
- Gelfand SA. *Hearing - an introduction to psychological and physiological acoustics*, 4th edition revised and expanded. Fourth edition, Taylor & Francis, 2005.
- Gilbert G, Micheyl C. Influence of competing multi-talker babble on frequency-importance functions for speech measured using a correlational approach. *Acta Acustica united with Acustica* 2005;91:145-54.

- Glasberg BR, Moore BC. Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *J Acoust Soc Am*. 1986;79:1020-33.
- Grant KW, Braida LD. Evaluating the articulation index for auditory-visual input. *J Acoust Soc Am*. 1991;89:2952-60.
- Hill FJ, McRae LP, McClellan RP. Speech recognition as a function of channel capacity in a discrete set of channels. *J Acoust Soc Am*. 1968;44:13-8.
- Hogan CA, Turner CW. High-frequency audibility: benefits for hearing-impaired listeners. *J Acoust Soc Am*. 1998;104:432-41.
- Hornsby BW, Ricketts TA. The effects of hearing loss on the contribution of high- and low-frequency speech information to speech understanding. *J Acoust Soc Am*. 2003;113:1706-17.
- Hornsby BW, Ricketts TA. The effects of hearing loss on the contribution of high- and low-frequency speech information to speech understanding. II. Sloping hearing loss. *J Acoust Soc Am*. 2006;119:1752-63.
- Horwitz AR, Ahlstrom JB, Dubno JR. Factors affecting the benefits of high-frequency amplification. *J Speech Lang Hear Res*. 2008;51:798-813.
- Houtgast T, Steeneken HJM. A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. *J Acoust Soc Am*. 1985;77:1069-77.
- Lieberman MC, Dodds LW. Single-neuron labeling and chronic cochlear pathology. III. Stereocilia damage and alterations of threshold tuning curves. *Hear Res*. 1984;16:55-74.
- Lippmann RP. Accurate consonant perception without mid-frequency speech energy. *IEEE Trans Speech Audio Process*. 1996;4:66-9.
- Macmillan NA. Beyond the categorical/continuous distinction: A psychophysical approach to processing modes. In S. Harnad (ed.). *Categorical perception: The groundwork of cognition*, New York: Cambridge University Press, 1987:53-88.
- Magnusson L. Predicting the speech recognition performance of elderly individuals with sensorineural hearing impairment. A procedure based on the Speech Intelligibility Index. *Scand Audiol*. 1996;25:215-22.
- Massaro DW. *Speech perception by ear and eye: a paradigm for psychological inquiry*. Earlbaum, Hinsdale, NJ, 1987.
- Moore BCJ. *Cochlear hearing loss - Physiological, Psychological and Technical Issues*. Second edition, Wiley, 2007.
- Moore BCJ. Dead Regions in the Cochlea: Diagnosis, Perceptual Consequences, and Implications for the Fitting of Hearing Aids. *Trends Amplif*. 2001;5:1-34.
- Moore BCJ, Vinay SN. Enhanced discrimination of low-frequency sounds for subjects with high-frequency dead regions. *Brain* 2009; 132:524-36.
- Pavlovic CV. Use of the articulation index for assessing residual auditory function in listeners with sensorineural hearing impairment. *J Acoust Soc Am*. 1984;75:1253-8.

- Pavlovic CV, Studebaker GA, Sherbecoe RL. An articulation index based procedure for predicting the speech recognition performance of hearing-impaired individuals. *J Acoust Soc Am*. 1986;80:50-7.
- Ronan D, Dix AK, Shah P, Braida LD. Integration across frequency bands for consonant identification. *J Acoust Soc Am*. 2004;116:1749-62.
- Sherbecoe RL, Studebaker GA. Audibility-index predictions of normal-hearing and hearing-impaired listeners' performance on the connected speech test. *Ear Hear*. 2003;24:71-88.
- Smootenburg GF. Speech reception in quiet and in noisy conditions by individuals with noise-induced hearing loss in relation to their tone audiogram. *J Acoust Soc Am*. 1992;91:421-37.
- Studebaker GA, Pavlovic CV, Sherbecoe RL. A frequency importance function for continuous discourse. *J Acoust Soc Am*. 1987;81:1130-8.
- Studebaker GA, Sherbecoe RL, McDaniel DM, Gwaltney CA. Monosyllabic word recognition at higher-than-normal speech and noise levels. *J Acoust Soc Am*. 1999;105:2431-44.
- Turner CW, Cummings KJ. Speech audibility for listeners with high-frequency hearing loss. *Am J Audiol*. 1999;8:47-56.
- Turner CW, Kwon BJ, Tanaka C, Knapp J, Hubartt JL, Doherty KA. Frequency-weighting functions for broadband speech as estimated by a correlational method. *J Acoust Soc Am*. 1998;104:1580-5.
- Vestergaard MD. Dead regions in the cochlea: implications for speech recognition and applicability of articulation index theory. *Int J Audiol*. 2003;42:249-261.
- Vickers DA, Moore BC, Baer T. Effects of low-pass filtering on the intelligibility of speech in quiet for people with and without dead regions at high frequencies. *J Acoust Soc Am*. 2001;110:1164-75.
- Warren RM, Riener KR, Bashford JA Jr, Brubaker BS. Spectral redundancy: intelligibility of sentences heard through narrow spectral slits. *Percept Psychophys*. 1995;57: 175-82.
- Zeng FG, Grant G, Niparko J, Galvin J, Shannon R, Opie J, Segel P. Speech dynamic range and its effect on cochlear implant performance. *J Acoust Soc Am*. 2002;111:377-86.
- Zwicker E, Schorn K. Psychoacoustical tuning curves in audiology. *Audiology*. 1978;17:120-40.

I.4. Les systèmes de réhabilitation de la fonction auditive

Le rôle d'un appareillage auditif est de restaurer au mieux la fonction auditive, à savoir l'audibilité, mais surtout l'intelligibilité. Nous avons vu précédemment que les deux principales conséquences d'une perte auditive étaient des distorsions d'intensité et de fréquence. L'aide auditive, par amplification, pourra redonner au malentendant la perception des petits sons, c'est-à-dire l'audibilité ; cependant pour les pertes les plus importantes, elle ne pourra certainement jamais palier des dommages tels que l'élargissement des filtres auditifs. D'une part ces filtres sont élargis du fait que les CCE ont été détruites par la pathologie, et ce phénomène n'est pas réversible. D'autre part le phénomène d'élargissement des filtres auditifs est d'autant plus important que l'intensité de stimulation est forte. Or pour que les sons soient audibles, il faut que l'aide auditive les amplifie. Le malentendant est donc toujours stimulé avec des intensités relativement fortes, ceci au détriment de la sélectivité fréquentielle.

Comme son nom l'indique clairement, l'aide auditive reste une « aide » pour la compréhension, mais en aucun cas elle ne pourra redonner une audition normale, aussi perfectionnée soit-elle. En ayant conscience de ces limitations, nous adapterons l'appareil avec pour objectifs principaux :

- amplifier sélectivement les sons afin de rendre perceptibles les sons de faible intensité et d'équilibrer les seuils de perception des différents sons (= palier la perte d'audibilité) ;
- limiter l'amplification des sons forts afin que les sons forts en entrée ne soient pas agressifs, et pour qu'ils soient mieux tolérés par le patient malentendant (= palier le recrutement de sonie) ;
- adapter la dynamique des sons conversationnels (la zone [30–70dB]) dans la dynamique résiduelle du sujet appareillé (figure 9).

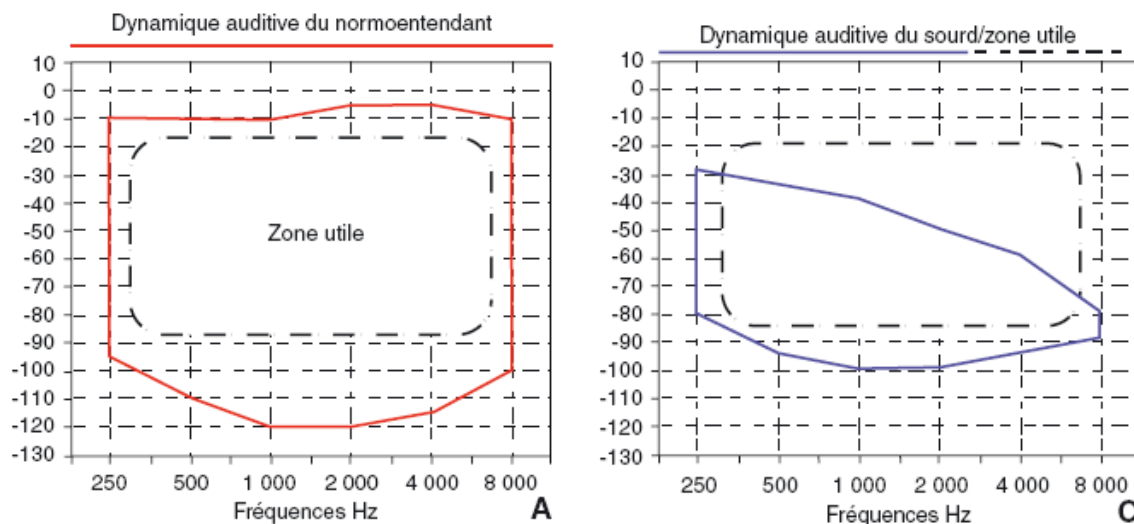


Figure 9. Partie gauche : zone utile des sons conversationnels versus dynamique auditive du normo-entendant. Partie droite : zone utile des sons conversationnels versus dynamique auditive du malentendant. D'après Lina-Granade et al., 2010

I.4.1. L'aide auditive

Principe

L'appareillage auditif par voie aérienne est destiné à la correction des surdités de perception légères à profondes. Le microphone de l'appareil va capter les sons de l'environnement et les amplifier sélectivement en leur accordant plus ou moins d'importance. Ensuite le signal passe dans un écouteur qui transmet le signal amplifié dans le conduit auditif du malentendant.

L'aide auditive peut se présenter classiquement (cf figure 10 : Lina-Granade et al. 2010) :

- sous la forme d'un intra-auriculaire, dans ce cas l'appareil est placé dans le conduit auditif externe ;
- sous la forme d'un contour ouvert, l'appareil est alors placé derrière le pavillon et le son est conduit jusqu'au tympan par un micro-tube, ce système a pour avantage de limiter l'effet d'occlusion de l'oreille ;
- sous la forme d'un contour à écouteur déporté, dans ce cas l'appareil est porté derrière le pavillon mais l'écouteur est placé dans le conduit auditif externe, il a l'avantage d'être beaucoup plus proche du tympan ;
- sous la forme d'un contour d'oreille classique ou surpuissant, qui permet de délivrer des puissances beaucoup plus importantes.

Les appareils auditifs ont toutefois des limitations en particulier lorsque la perte auditive devient trop importante. D'un point de vue technique le problème majeur va venir du phénomène de larsen qui se produit lorsque l'amplification dans les aigus est trop forte. D'un point de vue perceptif, et en particulier si le patient présente des zones inertes cochléaires (> 90 dB de perte), une amplification risque de créer des distorsions au point d'interférer avec les sons perçus par les zones cochléaires encore saines.

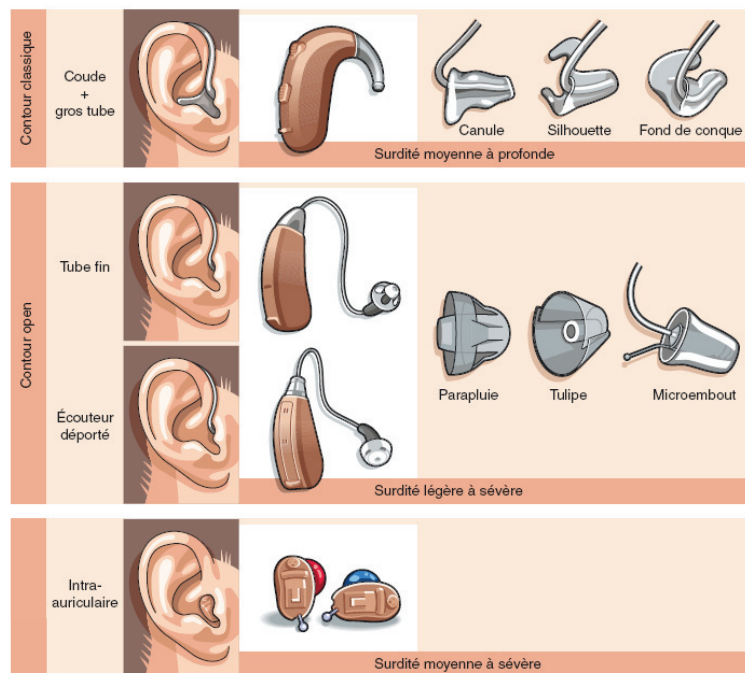


Figure 10. Les différents types d'aides auditives à conduction aérienne. D'après Lina-Granade et al., 2010.

I.4.2. L'implant cochléaire

Principe

L'implantation cochléaire est proposée aux malentendants ayant une surdité profonde à totale et depuis quelques années l'implant peut être proposé aux personnes présentant une perte auditive sévère. Pour ces malentendants l'appareillage par amplification n'est plus suffisant pour fournir une bonne compréhension de la parole, même avec des contours surpuissants ; lorsque l'intelligibilité en mots dissyllabiques devient inférieure à 50%, à 65 dB HL avec des aides auditives numériques bien réglées, on peut considérer que l'implant cochléaire fournira un bénéfice supérieur à l'aide auditive (Truy et al., 2009 ; Eshraghi et al., 2009).

Le principe de l'implant cochléaire est de réhabiliter la surdité par une stimulation électrique des voies auditives. L'objectif étant de fournir au cerveau un message qui reproduit au mieux le codage des sons dans l'audition naturelle, le signal délivré par l'implant doit avoir un codage adapté en fréquence et en intensité. Le patient implanté devra ensuite faire un gros travail de rééducation auditive et orthophonique, pour apprendre à décoder les sons nouvellement perçus.

L'objectif de l'implant est de restituer au mieux l'information temporelle (associée à la cadence de stimulation) et l'information fréquentielle (associée au nombre de canaux) des sons de parole prioritairement ainsi que des autres sons environnants.

Depuis les 2 ou 3 dernières décennies, l'implant cochléaire s'est considérablement amélioré. Aujourd'hui il n'est pas rare de voir des patients implantés qui téléphonent ou qui arrivent à suivre une conversation dans le bruit, ce qui n'était pas imaginable 30 ans auparavant. Mais cette évolution des performances n'est pas purement attribuable à la technologie. En parallèle, il y a eu une évolution des indications qui a permis de mieux sélectionner les candidats à l'implant.

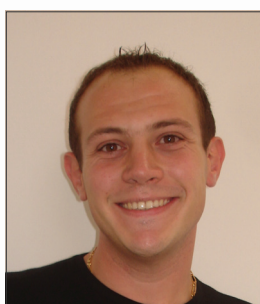
Au début de l'implant cochléaire, seuls les patients qui n'avaient aucune perception auditive pouvaient être candidats à l'implantation. Depuis le milieu des années 80, les indications audiométriques se sont relaxées et les centres d'implantation ont tendance à implanter des malentendants qui ont de plus en plus d'audition résiduelle dans les basses fréquences. Nous pouvons le constater sur une figure de Tyler et al. (2002) (cf figure 21 de l'article 3) qui montre les seuils préopératoires à 250 Hz de tous les patients qui ont été implantés à l'Université d'Iowa entre 1986 (seuils à 120 dB HL) et 1999 (seuils entre 30 et 85 dB HL).

Face à des patients implantés cochléaires qui arrivent à suivre une conversation (même au téléphone) et à des patients sourds sévères ou profonds qui sont constamment en difficulté dans ces mêmes situations de la vie quotidienne, on peut s'interroger sur la frontière entre l'appareillage acoustique et l'implant. Les patients implantés sont-ils réellement plus performants que les patients réhabilités acoustiquement ou bien est-ce une idée reçue ? Quel système recommander alors pour les patients qui sont à la limite ?

I.5. Evaluation des sourds réhabilités : De l'aide auditive à l'implant cochléaire

(Article publié dans ORL autrement : Seldran F, Gallégo S, Bigeard T, Thai-Van H, Truy E.)

■ Evaluation des sourds réhabilités de l'aide auditive à l'implant cochléaire bilatéral



Seldran Fabien

Doctorant en
Biomédical
Audioprothésiste DE



Gallégo Stéphane

Docteur en
Biomédical
Audioprothésiste DE,

**Thiphaine Bigeard
Thai-Van Hung
Eric Truy**

Introduction

Avec 5 millions de sourds et malentendants en France, la surdité est un problème qui touche de nombreuses personnes avec des conséquences importantes sur le plan social, économique,

Le champ des aides techniques pour ces patients ne cesse de se développer : les implants cochléaires se sont beaucoup répandus ces dernières années (miniaturisation, stimulation mixte électroacoustique) tandis que les prothèses auditives gagnent en performance grâce aux progrès de la technologie (compression fréquentielle). Compte tenu de ces évolutions, il paraît nécessaire de s'intéresser aux performances de ces différentes technologies et aux services rendus aux patients. L'évolution conjointe des techniques de la prothèse auditive et de l'implant cochléaire a permis d'élargir les indications et de mieux satisfaire les patients, mais sommes nous sûrs de connaître les performances de nos patients ?

Pour tout un chacun, les patients porteurs d'un implant cochléaire seraient plus performants en compréhension dans le silence et dans le bruit que les sujets sourds profonds porteurs d'une prothèse auditive conventionnelle, aussi puissante soit-elle, et à déficit auditif équivalent. Qu'en est-il réellement ? C'est ce que notre étude va s'efforcer d'évaluer afin de situer les performances des différentes populations d'implantés par rapport aux différentes populations de sourds appareillés.

Les aides auditives conventionnelles :

La surdité peut être en partie corrigée par l'utilisation d'aides auditives. L'objectif de l'appareillage est de compenser autant que faire se peut les déficits du système auditif du patient en améliorant les seuils auditifs ainsi que l'intelligibilité des sons de parole. Depuis quelques années déjà, de nombreux systèmes ont été mis au point pour optimiser les performances notamment dans le bruit.

Parmi ces systèmes, on distingue :

- les microphones directionnels,
- les réducteurs de bruit,
- les traitements de la parole.

Le bénéfice apporté par une aide auditive dépend surtout du type et du degré de surdité. Afin d'améliorer l'adaptation, il serait intéressant de pouvoir prédire le bénéfice de l'appareillage à partir des données audiométriques.

Cependant, l'expérience montre que des audiogrammes identiques conduisent à des apports prothétiques différents, suggérant que d'autres facteurs interviennent dans l'efficacité de l'appareillage.

En fait, le patient non appareillé n'entend pas certains signaux sonores uniquement parce que son seuil de perception est supérieur au niveau sonore à détecter.

Une fois correctement appareillé, ce même sujet entend les signaux sonores mais peut ne pas les comprendre. Les performances sont d'autant moins bonnes que les conditions d'écoute sont dégradées. Plusieurs explications peuvent être évoquées.

Facteurs liés à la périphérie : étant donné que les cellules ciliées externes (CCE) sont beaucoup plus fragiles que les cellules ciliées internes, les surdités de perception s'accompagnent systématiquement d'une altération des CCE, ce qui provoque une dégradation de la sélectivité en fréquence. Or, l'intelligibilité de la parole, surtout dans le bruit, dépend de la sélectivité en fréquence. La compensation de la perte par une aide auditive n'améliore pas la sélectivité en fréquence et laisse donc souvent persister des difficultés de compréhension dans le bruit. Néanmoins, depuis quelques années, les aides auditives ont développé des systèmes de prétraitement tels que des réducteurs de bruit ou de gestion de microphone qui permettent en partie d'améliorer le rapport signal sur bruit.

Facteurs centraux : les capacités cognitives des sujets âgés sont en moyenne amoindries par rapport à celles des sujets jeunes ; l'adaptation et les résultats de l'appareillage sont donc dépendants de l'âge du sujet. Pour les personnes ayant un potentiel cognitif réduit, il est nécessaire d'utiliser les appareils réduisant le bruit de fond, qui aident à mieux extraire le signal. Malgré ces difficultés, il n'existe pas de limite d'âge pour tirer bénéfice d'un appareillage auditif.

Les progrès techniques des appareils auditifs ont été considérables au cours des quinze dernières années, permettant un élargissement des indications et une meilleure satisfaction des patients. Cependant, la dégradation du message sonore au niveau de la cochlée reste le principal des problèmes de l'appareillage auditif.

L'implant cochléaire

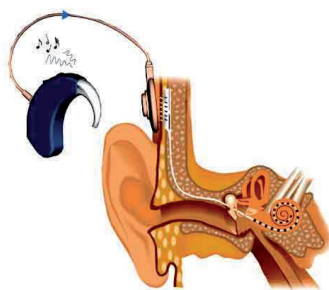


Figure 1 : Schéma de principe de l'implant cochléaire

L'implant cochléaire est un appareillage destiné aux personnes souffrant de surdité neurosensorielle profonde à totale non appareillables avec une aide auditive clas-

sique. Les indications actuelles de l'implant cochléaire considèrent qu'il faut une intelligibilité inférieure ou égale à 60% à 65 dB SPL avec une ACA moderne et bien réglée pour pouvoir bénéficier de l'implant. Il vise à induire des sensations auditives, en stimulant directement le nerf auditif (VIII) au moyen d'une vingtaine d'électrodes implantées chirurgicalement, et dont la séquence de stimulation issue du stimulus est réglée en s'inspirant du mode de fonctionnement du code neural. Malgré les limitations de la vingtaine d'électrode vis-à-vis des 3000 cellules ciliées internes d'une personne normo entendant, l'implant cochléaire est suffisant pour restituer la parole intelligiblement. L'implant cochléaire se substitue à la transduction cochléaire défectueuse chez le patient atteint de surdité neurosensorielle : en ce sens, il s'agit à proprement parler d'un organe sensoriel à part entière, ce qui lui vaut le surnom d'oreille bionique. L'efficacité de l'implant cochléaire résulte de l'exploitation des 2 types de codages des sons qui accompagnent la transduction. La transduction cochléaire se base sur l'organisation tonotopique de la cochlée (basses fréquences codées à l'apex, et hautes fréquences à la base) et sur la synchronisation du système auditif sur la périodicité du stimulus (phase-locking). L'implant cochléaire peut être indiqué pour tous types de surdités neurosensorielles profondes à totales liées à la non-fonctionnalité des cellules ciliées de la cochlée.

L'adaptation à l'implant est souvent facile. Dès la première semaine d'utilisation, le patient implanté peut acquérir une bonne perception du langage sans qu'une réhabilitation spécialisée soit indispensable. Cependant de bien meilleurs résultats sont souvent obtenus lorsque la personne suit des séances d'orthophonie et d'entraînement. Le bénéfice que la personne va retirer de son implant va beaucoup évoluer dans les premiers mois après l'implantation. Chez l'adulte, la perception de la parole juste après implantation est en effet à peu près de moitié ce qui sera finalement obtenu. La plus grande partie de l'amélioration se produit dans les neuf premiers mois après implantation, particulièrement s'il y a de l'entraînement auditif intensif. Cependant, en moyenne, la reconnaissance de la parole s'améliorera jusque dans les 18-30 mois après implantation chez l'adulte. Les implantés cochléaires présentent une grande variabilité interindividuelle dans leurs performances auditives, en particulier en terme de perception de la parole sans que l'on sache vraiment pourquoi. Certains patients pourront communiquer par téléphone tandis que d'autres sont seulement

capables de détecter des signaux sonores. Par ailleurs, il est supposé que la performance avec un implant est fortement associée avec les capacités de traitement auditif de la personne et de l'intégrité du système auditif central depuis le nerf auditif jusqu'au cortex. Il apparaît aussi qu'une courte durée de surdité profonde et la présence de reliquats auditifs soient deux facteurs prédictifs majeurs. D'autres facteurs tels que le nombre de cellules survivantes du ganglion spiral, l'étiologie de la surdité et le temps d'utilisation de l'implant sont aussi associés à des changements de performances auditives.

L'âge d'implantation est également un paramètre très important à prendre en compte. La réhabilitation d'une personne sourde post-linguale est, en général, plus simple que dans le cas de surdités pré-linguales. En effet, ces personnes peuvent avoir recours à leur mémoire auditive pour établir une relation entre les nouvelles sensations auditives et l'audition perdue. En 1991, les adultes sourds sévères à profonds pouvaient déjà atteindre des niveaux perceptifs moyens de mots de 14 à 50 % et de 40 à 90 % pour des phrases sans aide de la lecture labiale et en liste ouverte. Les performances actuelles obtenues avec des implants cochléaires sont très bonnes dans le silence mais aussi dans le bruit.

Lorsque l'oreille controlatérale possède quelques restes auditifs exploitables, il est généralement conseillé d'essayer en complément à l'implant cochléaire une aide auditive conventionnelle. Des études ont montré que cela améliorerait l'intelligibilité dans le silence et dans le bruit et surtout que cela rendait un équilibre spatial et une meilleure localisation.

Depuis une dizaine d'année des études ont aussi montré un bénéfice fonctionnel à l'implantation cochléaire bilatérale. Cependant des facteurs économiques et des risques liés à l'altération du système vestibulaire ont un peu freiné son développement.

L'implant électroacoustique (EAS)

Il est conçu pour des personnes présentant une surdité partielle. Cette surdité se définit comme une surdité neurosensorielle légère à modérée dans les fréquences graves, associée à une surdité profonde dans les fréquences aiguës. L'EAS est constitué de deux parties : d'une part un implant cochléaire, composé d'une électrode flexible, stimule électriquement les cellules nerveuses responsables du

traitement des fréquences aiguës. D'autre part, une amplification acoustique stimule les cellules nerveuses responsables du traitement des fréquences graves. L'association de la stimulation électrique à la stimulation acoustique permet au patient d'obtenir une audition de qualité qui exploite son audition résiduelle naturelle. L'implant EAS, comme l'implant cochléaire est destiné aux sujets appareillés en ACA dont les performances ne dépassent pas 60 % à 65 dB SPL. Il est indiqué aux sujets qui ont des restes auditifs exploitables sur les fréquences graves (250 à 750 Hz).

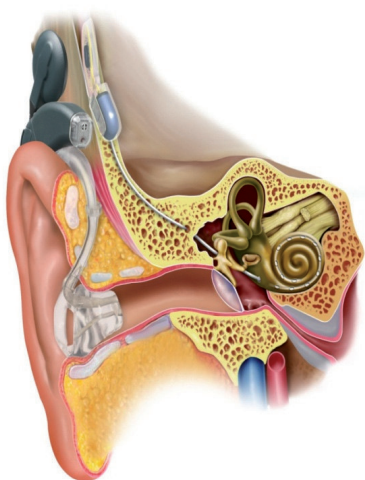


Figure 2 : Schéma de principe de l'implant électro-acoustique

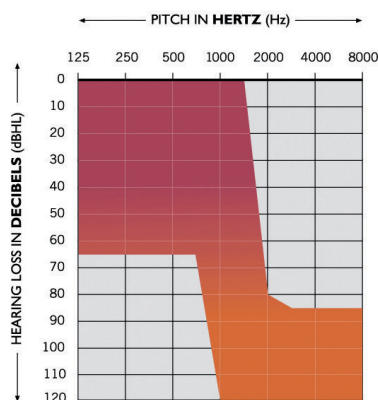


Figure 2 bis : Plage d'application

Problématique et hypothèse

Nous avons vu qu'il existe deux grandes catégories d'aides techniques pour les patients présentant différents types de surdités. Nous pouvons nous interroger sur l'aide apportée par chaque appareillage en fonction du degré de surdité. La littérature évoque une grande variabilité quant à la performance auditive des patients implantés. Des auteurs tel Blamey en 1996 ; Gallego et al, 1998, ont évoqué des facteurs pouvant être associés à des changements de performances auditives. L'objectif de cette étude est de situer les performances obtenues avec un implant cochléaire par rapport à celle obtenues avec des systèmes conventionnels (aides auditives) pour des patients adultes ayant différents types de surdité. Nous nous attendons à ce que les patients IC aient de meilleures performances que les patients sourds profonds et sévères dans le calme (1^{ère} hypothèse). La 2^{ème} hypothèse : ce peut être la même chose mais dans le bruit. En dissociant les types d'implantation, nous verrons si les scores sont les mêmes chez les implantés unilatéraux et les bilatéraux.

Populations étudiées

Une population témoins de 20 normo-entendants a servi de références.

La population sourde étudiée est composée d'une part de malentendants appareillés depuis plus d'1 an et d'autre part de patients implantés cochléaires depuis plus d'un an.

Tous les patients appareillés sont porteurs de deux prothèses auditives.

On distingue parmi eux :

- 15 surdités légères,
- 41 surdités moyennes,
- 15 surdités sévères ou profondes.

Pour les patients implantés, on distingue :

- 26 patients avec un implant cochléaire unilatéral seul,
- 13 patients avec un implant cochléaire avec une aide auditive conventionnelle controlatérale,
- 6 patients avec un implant électro-acoustique (EAS),
- 7 patients implantés cochléaires bilatéralement.

Protocole

L'objectif étant de situer les performances de la population implantée cochléaire à celles d'une population appareillée avec des prothèses auditives conventionnelles,

pour ce faire plusieurs tests ont été pratiqués.

Pour les patients appareillés (aca) :

- une anamnèse pour vérifier nos critères d'inclusion et une otoscopie afin d'écarter tout bouchon de cérumen qui fausserait les tests,
- une audiométrie tonale au casque (en conduction aérienne) avec recherche du seuil de perception par demi-octaves du 250Hz au 8000Hz pour chaque oreille avec des sons purs,
- un gain prothétique tonal avec leurs deux prothèses auditives en champ libre. Les fréquences testées vont du 250Hz au 6000Hz. On utilise un son volubé qui semble plus facile à identifier,
- une audiométrie vocale dans le silence oreilles nues avec des listes de mots dissyllabiques de Fournier (40 listes de 10 mots) numérisés avec une voix d'homme. Nous cherchons l'intensité minimum à laquelle le patient obtient 100% d'intelligibilité,
- une audiométrie vocale dans le silence avec leurs deux prothèses auditives. Nous utilisons toujours les listes de mots dissyllabiques de Fournier numérisés et recherchons le 100% d'intelligibilité à l'intensité la plus faible.

- Le test d'Elbaz

C'est un test de résistivité au bruit. C'est une audiométrie vocale dans le bruit. Il s'agit dans nos tests d'un bruit blanc qui est diffusé en même temps que la parole avec un rapport signal/bruit que l'on fait varier par pas de 3 dB. L'intensité de la parole est fixe durant le test. Sa valeur correspond à l'intensité du plateau d'intelligibilité obtenu lors de l'audiométrie vocale dans le silence + 10dB. Les tests vocaux sont effectués en champ libre d'abord oreilles nues puis oreilles appareillées. La parole et le bruit sont diffusés à travers deux haut-parleurs situés face au patient. Lors du test le bruit est progressivement augmenté jusqu'à atteindre 50% d'intelligibilité. Pour toutes les audiométries vocales et pour ne pas influencer les résultats par l'ordre de passage des listes de mots, les passations sont réalisées en ordre aléatoire.

Pour les patients implantés

Pour tous ces patients, on réalise lors de leur visite à l'hôpital HEH :

- un gain prothétique tonal en champ libre avec leurs appareils comme ils les

portent au quotidien. On utilise un son vobulé qui sort par deux haut-parleurs situés face à eux et l'on balaye les fréquences de 250Hz au 6000Hz,

- une audiométrie vocale dans le silence avec appareils avec des listes de mots dissyllabiques de Fournier comme pour les patients appareillés.

- une audiométrie vocale dans le bruit (bruit blanc) avec le test d'Elbaz et les listes de mots dissyllabiques de Fournier.

Les conditions de tests sont les mêmes pour les implantés, les patients appareillés en conventionnel, et les sujets témoins. On utilise la même méthode de comptage pour avoir le score des audiométries vocales c'est à dire chaque liste de Fournier contient 10 mots dissyllabiques, donc chaque syllabe juste compte pour 10% de réussite.

Les voyelles réalisées en champ libre nous permettent de définir l'intelligibilité du patient. On a pu avec les tests statistiques (régression non linéaire sigmoïdale) définir le plateau d'intelligibilité et le seuil à la moitié de l'intelligibilité plateau. Par ailleurs, le test d'Elbaz nous a permis de calculer le 50% d'intelligibilité dans le bruit (seuil en dB aussi déterminé par régression non linéaire sigmoïdale).

Pour l'analyse des résultats, des tests statistiques sont utilisés. Des tests anova permettent de comparer tous les paramètres entre eux afin de voir les corrélations possibles.

Résultats

Seuil tonal

Nous avons déterminé les seuils prothétiques en tonale de chaque groupes de patients en faisant la moyenne des seuils en champ libre au 500, 1000, 2000, 4000Hz. Les résultats obtenus avec Anova, nous ont classifiés du moins bon au meilleurs les différents groupes de patients avec pour chacun d'eux leur seuil tonal appareillé moyen en champ libre en dB HL ($F=35.8$, $p<0.0001$). (Fig : 3)

On constate que les seuils en tonale des patients implantés sont bons. Ils sont même meilleurs que ceux des patients atteints d'une surdité moyenne, sévère et profonde. L'écart entre l'implant unilatéral et les surdités sévère et profonde est très important avec 17dB de différence. Les

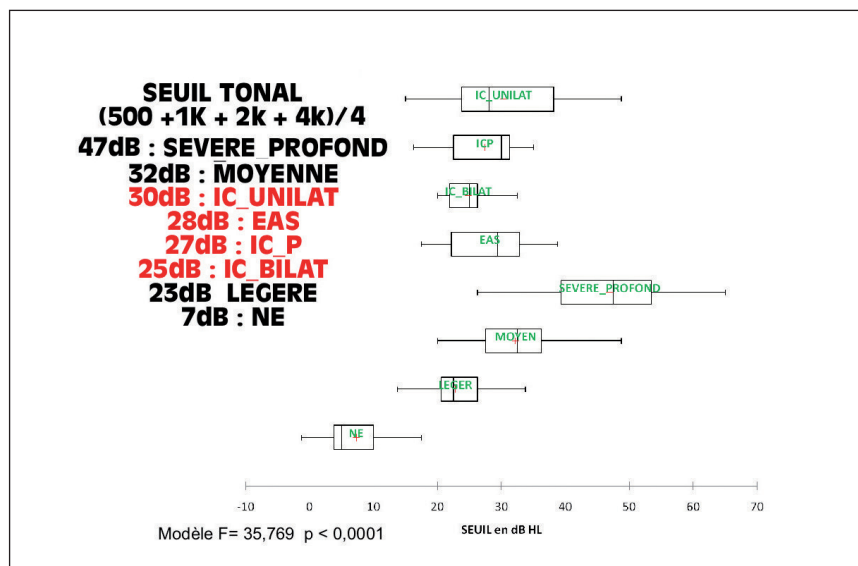


Figure 3 : Seuil Tonal moyen pour les différents groupes de sujets et classement du moins bon au meilleur groupe. Les groupes implantés sont encadrés par les groupes appareillés.

patients porteurs d'un implant unilatéral ont un seuil moins bon que les patients porteurs de deux implants ou d'un implant + prothèse. Les valeurs du seuil perceptif champ libre pour l'ensemble des groupes sont peu différentes entre elles (proche de 30 dB), excepté pour les surdités sévères et profondes. Il y a une certaine homogénéité dans les résultats. Les patients implantés sont meilleurs comparés aux surdités moyennes, sévères et profondes.

Intelligibilité :

Afin de comprendre notre méthode de régression non linéaire sigmoïdale, prenons l'exemple d'un patient implanté en unilatéral. (Fig : 4)

Ce schéma représente un exemple de résultats de l'audiométrie vocale appareillée en champ libre chez un sujet implanté cochléaire. Avec l'équation de la régression non-linéaire sigmoïdale, nous avons pu déterminer pour chaque sujet le plateau d'intelligibilité (ici 94,3 %) et le seuil à la moitié de l'intelligibilité plateau (ici 31,5 dB SPL).

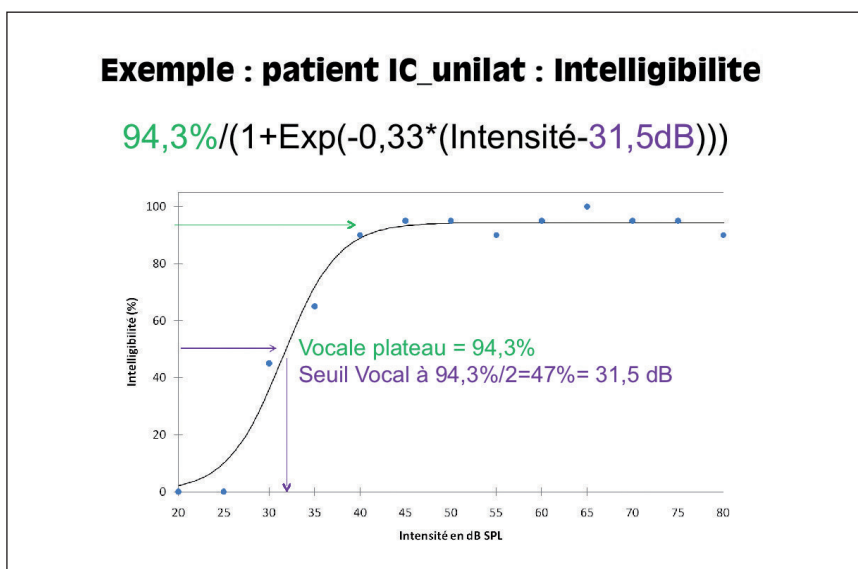


Figure 4 : Méthode permettant de mesurer l'intelligibilité plateau ainsi que le seuil vocal à la moitié de cette valeur. La fonction sigmoïdale permettant de modéliser les résultats permet de déterminer précisément le plateau et le seuil d'intelligibilité.

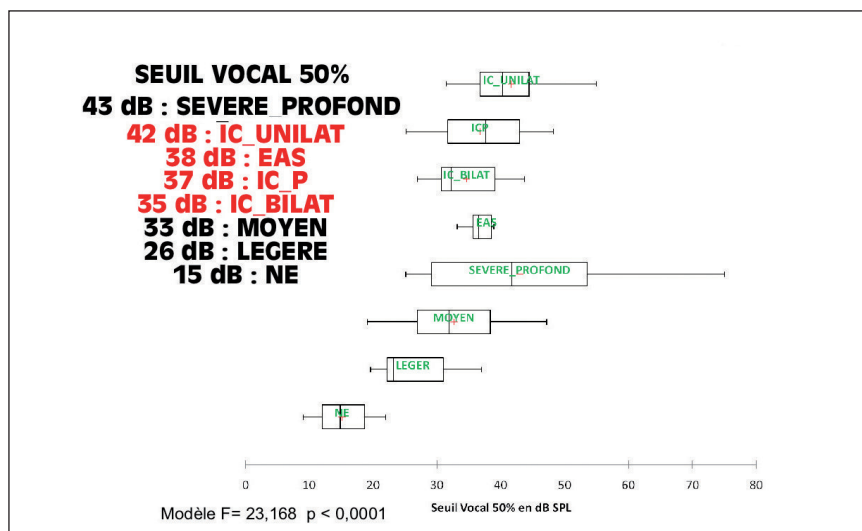


Figure 5 : Seuil vocal moyen à 50% du plateau pour les différents groupes de sujets et classement du moins bon au meilleur groupe. Seul le groupe appareillé sourd sévère-profond est moins bon que les groupes implantés.

Seuil vocal 50% du plateau

Les résultats obtenus avec Anova, nous ont classifiés du moins bon au meilleurs les différent groupes de patients avec pour chacun d'eux leur seuil d'intelligibilité appareillé moyen en champ libre en dB HL (F=23.2, p<0,0001). (Fig : 5)

Ces résultats représentent la valeur du seuil auditif en dB SPL de la moitié de l'intelligibilité plateau. Les implantés se situent entre les surdités moyennes et les surdités sévères-profondes. A noter que les patients en unilatéral sont bien en dessous des performances des autres groupes d'implantés. Ils ont presque le même seuil que les surdités sévères-profondes. La chronologie des

résultats souligne le fait que les implantés bilatéraux sont meilleurs que les implants + prothèse controlatérale qui sont meilleurs que les EAS.

Vocale plateau

L'audiométrie vocale plateau est la plus représentative des performances réelles des sujets car elle prend en compte tous les points testés. Elle n'a pas d'effet "listes" c'est-à-dire qu'au cours de l'audiométrie vocale, des listes de mots paraissent plus facile que d'autres au patient. Ce plateau est défini par test statistique (régression sigmoïdale de l'intelligibilité). Les résultats obtenus avec Anova, nous ont classifiés du moins bon au meilleurs les différents groupes de patients avec pour chacun d'eux leur intelligibilité moyenne plateau appareillé en champ libre en dB HL (F=13.8, p<0,0001). (Fig : 6)

Les résultats montrent que les implantés sont moins bons que les sourds appareillés excepté pour les EAS. Les EAS ont un score de 89,3 % qui est proche du score des surdités sévères-profondes (89,6 %). L'implant unilatéral est toujours le plus en difficulté avec un score de 79,8 %. L'implantation bilatérale améliore les performances. Les scores des surdités légères et moyennes sont excellents. Quand on regarde le schéma représentant les patients, on constate une disparité des résultats pour les implantés. Les performances sont très hétérogènes. Elles le sont davantage chez les implantés unilatéraux que chez les implants + prothèse ou implants bilatéraux. Sur les 52 implantés seuls 3 ont un plateau à 100%. Pour les 6 EAS, la valeur moyenne du plateau et la disparité des résultats sont identiques à ce qui est observé chez le groupe de surdités sévères-profondes, mais le faible nombre de patients ne permet pas d'en faire une généralisation.

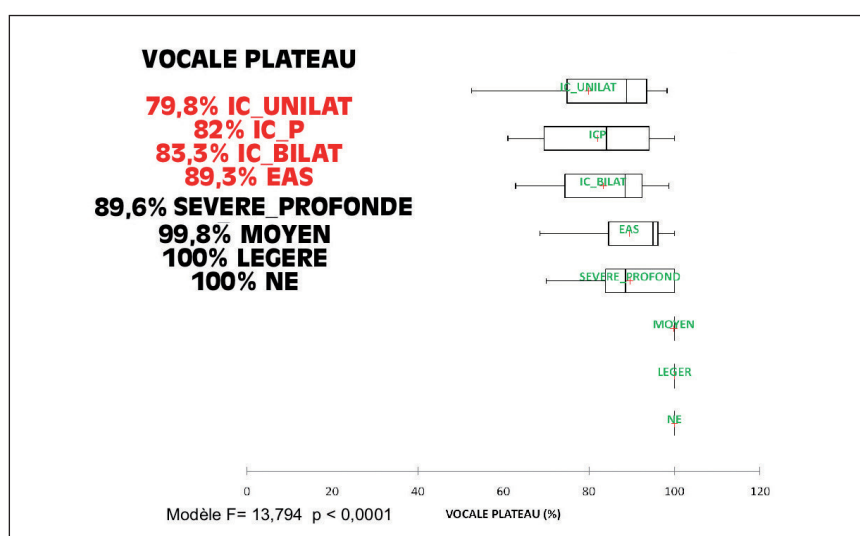


Figure 6 : Intelligibilité moyenne plateau pour les différents systèmes et classement du moins bon au meilleur groupe. Les groupes implantés sont moins bons que les groupes appareillés.

Exemple : patient IC_unilat :

Elbaz $99,0\%/(1+\text{Exp}(-0,34*(S/B -12,2\text{dB})))$

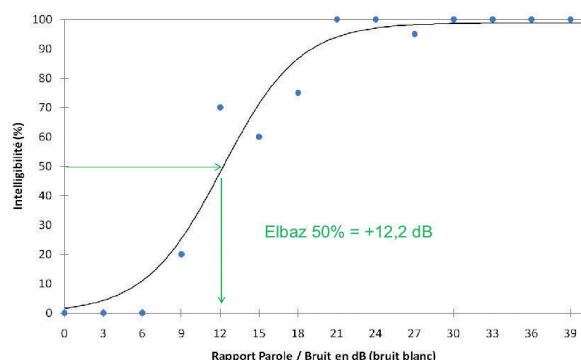


Figure 7 : Méthode permettant de mesurer le seuil à 50 % du test d'Elbaz.
La fonction sigmoïdale permettant de modéliser les résultats permet de déterminer précisément le rapport signal/bruit où l'intelligibilité est de la moitié du plateau (+12.2 dB sur cet exemple).

Test dans le bruit d'Elbaz :

Prenons l'exemple d'un patient implanté en unilatéral. (Fig : 7)

Comme pour l'intelligibilité dans le silence, les tests statistiques ont permis de tracer une régression sigmoïdale afin de calculer avec précision le seuil d'intelligibilité à 50% de la vocale dans le bruit.

Sur ce graphique, on remarque que ce patient implanté cochléaire unilatéralement n'est pas performant dans le bruit. Le rapport S/B reste positif au 50%. Dans ce cas, le 50% d'intelligibilité dans le bruit est de +12.2dB (c'est à dire qu'il faut que la parole soit plus forte de 12,2dB par rapport au bruit).

Test Elbaz-50(S/B) :

Les résultats obtenus avec Anova, nous ont classifiés du moins bon au meilleur les différents groupes de patients avec pour chacun d'eux leur rapport parole sur bruit en dB du seuil d'intelligibilité dans le bruit ($F=33.4$, $p<0,0001$). (Fig : 8)

Les résultats trouvés pour chaque groupe de patients soulignent la mauvaise performance de tous les groupes de patients implantés. En effet l'écart est significatif dès lors que le rapport S/B reste positif pour les implantés alors qu'il est négatif pour les autres. Une exception toutefois : les EAS qui ont un rapport S/B négatif à -1dB. Ils sont malgré tout moins bons que

les surdités sévères-profondes. Comme précédemment, les unilatéraux sont les plus mauvais parmi les implantés. Il y a une grande disparité entre les sourds appareillés et les patients implantés. Les implantés ont une plus grande difficulté à comprendre dans le bruit. Toutes les surdités appareillées ont un rapport S/B moyen négatif c'est à dire qu'ils sont encore performants lorsque le bruit est supérieur à la parole.

Discussion

Au niveau des seuils en audiométrie tonale champ libre, on remarque que les patients porteurs d'un implant sont très bons. Ils se situent entre les surdités légères et les surdités moyennes.

Les implantés en unilatéral sont un peu moins bons que des implantés bilatéraux (5dB de différence) et implant + prothèse (3 dB de différence).

Les seuils pour les groupes implant + prothèse et implant bilatéral sont légèrement meilleurs que les EAS et les unilatéraux mais pas de façon significative.

Pour les valeurs du seuil au 50 % de la vocale, on retrouve une même homogénéité. Les implantés se situent toujours bien mais un peu en retrait par rapport au seuil tonal : ils sont entre les surdités moyennes et les surdités sévères-profondes.

Malgré des seuils de perception similaires et des seuils de vocale à 50% du plateau proches, la compréhension vocale des patients implantés est bien en deçà de celle des sourds appareillés. On peut affirmer que les patients implantés cochléaires ont une très bonne détection des sons mais cela ne leur suffit pas pour avoir une bonne compréhension dans le silence ou dans le bruit. Lorsque l'on regarde la vocale plateau, les implantés sont plus mauvais que les sourds appareillés avec prothèses. Les EAS cependant sont davantage performants que les autres implantés.

Dans l'étude, ils sont les patients implantés les plus performants dans le bruit, mais leur faible nombre ne nous permet pas de généraliser cette performance par rapport aux autres groupes. Ils restent toujours moins bons que les sourds sévères et profonds.

Au niveau de la "vocale plateau", les EAS ont un pourcentage d'intelligibilité très proche des surdités sévères et profondes (seulement 1% de différence).

Avec ces tests, nous avons constaté que les patients ayant deux appareils sont

TEST ELBAZ-50 (S/B)

5,5dB IC_UNILAT
3,1dB IC_P
2,8dB IC_BILAT
-1,0dB EAS

-2,3dB SEVERE_PROFOND
-9,8dB MOYENNE
-11,4dB LEGERE
-15,1dB NE

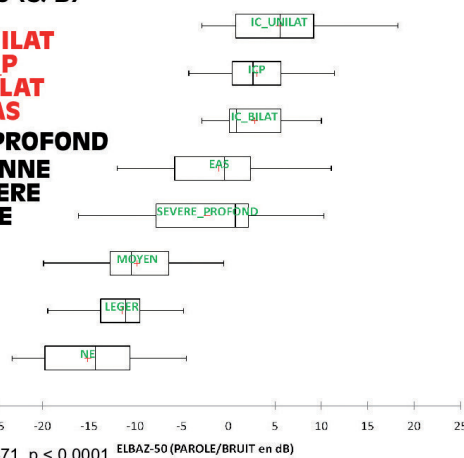


Figure 8 : Rapport signal bruit moyen pour obtenir 50% du plateau pour les différents systèmes et classement du moins bon au meilleur groupe.
Les groupes implantés sont moins bons que les groupes appareillés.

toujours meilleurs en vocale dans le silence et dans le bruit que les patients avec un implant seul.

L'implantation bilatérale, séquentielle ou simultanée permet d'améliorer la fonction binaurale : meilleure discrimination en condition acoustique défavorable, augmentation de la sonie en outre. Des études faites notamment par Medel ont montré que l'implantation bilatérale améliorerait nettement la compréhension dans le bruit. L'audition binaurale (qu'elle soit avec 2 implants ou IC+ prothèse) permet une meilleure perception de la parole dans le bruit. Toutefois une étude en 2006 précise que la compréhension du langage dans le bruit s'améliore avec une amplification binaurale, mais l'acceptation du bruit ne dépend pas de l'amplification mono ou binaurale pour la plus grande part des sujets (Melinda C. Freyaldenhoven en 2006).

Des études ont montré que le choix du type de bruit ne semble pas influencer sur l'intelligibilité de la parole (P.A Howard-Jones, S.Rosen 1993). De plus la préférence pour un bruit de fond semble indépendante de son acceptation par le patient. Le patient pourrait manquer d'aptitude à évaluer sa capacité à accepter les bruits de fond (Melinda C. Freyaldenhoven en 2006).

Dans la littérature, une étude précise que les variations physiologiques qui proviennent des régions plus centrales du système auditif peuvent intervenir dans l'acceptation du bruit de fond.

Cette étude avait pour but de comparer les performances entre patients IC et patients porteurs de prothèses. En les comparant aux surdités appareillées avec des prothèses auditives conventionnelles, on s'est aperçu qu'il y avait encore un écart très important dans les résultats en faveur des patients porteurs de prothèses pour ce qui est de la compréhension dans le bruit.

Il aurait été intéressant de tenir compte de la rééducation orthophonique dans notre étude. Il est certain que c'est un critère important qui influe sur la performance du patient quel qu'il soit. Il en est de même pour les réglages qui n'ont pas été pris en compte.

Par ailleurs des études faites à Montpellier précisent qu'un patient implanté depuis plus de 15 ans obtient de meilleurs résultats en compréhension avec une vitesse d'élocution plus lente. Il serait probablement plus objectif d'en tenir compte pour pouvoir comparer cette population implantée à celle appareillée avec des aides auditives conventionnelles.

Bien que l'adaptation à l'implant soit souvent facile (dès les premiers mois le patient implanté peut acquérir une bonne perception du langage), il semble qu'en situation de tests comme nous avons procédé lors de notre expérimentation, leurs performances sont peu satisfaisantes. En effet, cette étude nous montre que les patients implantés sont beaucoup plus en difficulté pour comprendre que les patients appareillés. Cependant, nous ne devons pas oublier de préciser qu'il existe une grande variabilité interindividuelle dans les performances de ces patients.

Comme nous l'avons vu, les patients porteurs d'un implant EAS semblent se détacher des autres groupes d'implantés, notamment au regard de leurs performances dans le bruit. Toutefois, avant de pouvoir confirmer cette tendance, il faudrait recruter un nombre beaucoup plus important de patients avec EAS.

Actuellement, les indications pour l'implant concernent les patients déficients auditifs qui ne tirent pas un bénéfice suffisant de la prothèse auditive en terme de compréhension de la parole. Or les critères audiologiques ont beaucoup évolué au cours de ces dernières années. L'indication, d'abord limitée à la surdité profonde de type III, a été étendue à toutes les surdités profondes. Aujourd'hui, les audiologistes s'interrogent même sur l'implantation possible de surdités sévères.

Il est certain que des avancées futures sont encore attendues tant dans le domaine de la prothèse auditive que dans celui de l'implant cochléaire. Aujourd'hui déjà se pose la question de l'implantation bilatérale (faut-il la faire en simultanée ou séquentielle). L'implantation bilatérale doit-elle devenir un standard comme l'appareillage stéréophonique avec des prothèses conventionnelles est devenu un standard ?

Il nous faudra donc affiner le plus possible les tests (de résistivité dans le bruit en particulier) pour faire le meilleur choix entre prothèses et implants (uni ou bilatéral, EAS ou non...) et tirer ainsi le plus profit de ces progrès technologiques. Il faut comprendre les enjeux de chaque type d'appareillage (coût économique, double chirurgie pour l'implant, coût pour le patient...).

En conclusion générale, même si les groupes de sujets implantés nous font croire de par un audiogramme tonal champ libre très valorisant (de 25 à 30 dB) que ce type de réhabilitation est meilleur que l'appareillage conventionnel (seuil de 47dB pour le groupe sévère-profond) ; les performances de compréhension de la parole dans le silence et dans le bruit situent l'implant

cochléaire au dernier rang. Ils sont donc à leur place ; c'est-à-dire une solution lorsque l'aide auditive conventionnelle n'apporte plus suffisamment. ■

Bibliographie

Blamey P., Arndt P., Bergeron F., Bredberg G., Brimacombe J., Facer G., Larky J., Lindström B., Nedzelski J., Peterson A., Shipp D., Staller S., Whitford L. (1996) Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants. *Audiol. Neurotol.* 1(5), 293-306.

Freyaldenhoven MC, Plyler PN, Thelin JW, Burchfield SB. (2006) Acceptance of noise with monaural and binaural amplification. *J. Am. Acad. Audiol.* 17(9), 659-66.

Freyaldenhoven MC, Smiley DF, Muenchen RA, Konrad TN. (2006) Acceptable noise level: reliability measures and comparison to preference for background sounds. *J. Am. Acad. Audiol.* 17(9), 640-8.

Gallégo S., Frachet B., Micheyl C., Truy E., Collet L. (1998) Cochlear implant performance and electrically-evoked auditory brainstem response characteristics. *Electroenceph. Clin. Neurophysiol.* 108, 521-525.

Howard-Jones P.A., Rosen S. (1993) Unmodulated glimpsing in "checkerboard" noise. *J. Acoust. Soc. Am.* 93(5), 2915-22.

Nous avons mis en évidence que malgré des audiogrammes meilleurs, les patients implantés cochléaires éprouvent beaucoup plus de difficultés que les patients réhabilités acoustiquement, pour la compréhension de la parole en particulier en situation de bruit concurrent. Comment expliquer cela ?

L'explication qui vient d'abord à l'esprit est la différence de modalité. Dans le cas de l'appareillage conventionnel, la stimulation est acoustique et la modalité est naturelle.

Pour l'implant, la modalité n'est pas naturelle et une réattribution tonotopique est nécessaire, ce qui va demander un gros travail de rééducation. Au-delà de cette explication, l'implanté cochléaire a des déficits perceptifs beaucoup plus importants que le sourd réhabilité acoustiquement. Les implantés ont d'une part un appauvrissement du codage fréquentiel mais aussi un appauvrissement du codage en intensité – rappelons que la dynamique électrique est de l'ordre de 10 à 15 dB.

Aussi, les performances avec l'implant sont très influencées par les caractéristiques propres au patient tels que l'âge d'apparition de la surdité, la durée de surdité pré-implantation, la proportion de vie sourde ou encore le degré d'audition résiduelle.

Dans la section suivante nous allons, par une brève revue de littérature, essayer de comprendre l'influence de ces différents facteurs sur les performances de l'implanté cochléaire.

I.6. Implant cochléaire et intelligibilité : Comment expliquer la disparité des performances ?

Implant cochléaire et intelligibilité : Comment expliquer la disparité des performances ?

(Article soumis aux Cahiers de l'audition)

Fabien Seldran (1,2), Doctorant en Ingénierie Biomédicale, audioprothésiste.

Lionel Collet (1), Professeur des universités-Praticien Hospitalier, Chef du service d'audiologie et d'explorations oro-faciales des Hospices Civils de Lyon, Directeur de l'école d'Audioprothèse de Lyon.

Stéphane Gallégo (1,3), Docteur en Ingénierie Biomédicale, audioprothésiste.

(1) Hôpital Edouard Herriot, Pavillon U, service d'audiologie et d'explorations oro-faciales, Lyon

(2) Vibant Med-El, Sophia-Antipolis

(3) Audition Conseil, Lyon

L'implant cochléaire (IC) s'est considérablement répandu ces dernières années, notamment grâce à la technologie numérique et à l'avancée des techniques de traitement du signal. A ses débuts l'IC n'était qu'une aide à la lecture labiale tandis qu'aujourd'hui les patients implantés parviennent généralement à comprendre dans les situations d'écoute facile et difficile. Cependant à l'échelle individuelle, nous constatons que certains implantés ont des scores de compréhension très satisfaisants et parviennent même à téléphoner avec leur implant, alors que d'autres patients éprouvent beaucoup de difficultés pour comprendre, malgré plusieurs années d'utilisation de l'implant et malgré une rééducation orthophonique intensive. Cette variabilité des performances dépend de nombreux facteurs. Certains de ces facteurs sont inhérents à l'implant, comme par exemple la génération de l'implant ou la stratégie de codage (Zeng, 2004 ; Zeng et al., 2008 : figure 20.b. et c.). D'autres facteurs sont inhérents au patient lui-même et à ses caractéristiques. Parmi ces facteurs, nous trouvons l'âge d'apparition de la surdité, la durée de surdité ou encore le degré d'audition résiduelle avant l'implantation. L'identification de cet ensemble de facteurs pronostiques et prédictifs a permis en particulier de mieux définir les indications et de mieux cibler les candidats à l'implantation.

Dans ce document, nous allons nous intéresser dans un premier temps aux travaux qui ont examiné les relations entre mesures objectives et perceptives chez ces patients. Par la suite, nous allons présenter les principaux articles scientifiques qui ont mis en évidence l'influence des différents paramètres décrits ci avant sur les performances post-implantation chez l'adulte devenu sourd.

Etat de la littérature

Récemment, Wilson et Dorman (2008) dans une revue de littérature, ont présenté les performances de 55 sujets implantés cochléaires (données recueillies dans l'étude de Helms et al., en 1997 actualisées par Wilson en 2006) utilisateurs du système Combi 40 de Med-El avec la stratégie CIS. Les sujets étaient soumis à des tests de compréhension de phrases et de mots monosyllabiques et les auteurs ont mesuré l'évolution des performances entre 1 et 24 mois post-activation (figure 1). Les scores d'intelligibilité atteignent un plateau au bout de 12 mois. Pour la compréhension des phrases qui est un test assez facile, à 24 mois, 46 des 55 sujets obtiennent des scores supérieurs à 80%. Pour la compréhension des mots monosyllabiques qui est un test plus difficile, les performances chutent à une moyenne de 55%, et seulement 9 patients testés obtiennent des scores supérieurs à 80%. Nous observons en particulier que la variabilité des résultats est beaucoup plus grande, avec certains patients qui obtiennent des scores proches de 100% et d'autres sujets qui stagnent autour de 10% d'intelligibilité. Du fait que les 55 patients utilisent tous le même système et la même stratégie de codage, cette variabilité des résultats ne peut pas être expliquée par des différences de technologies.

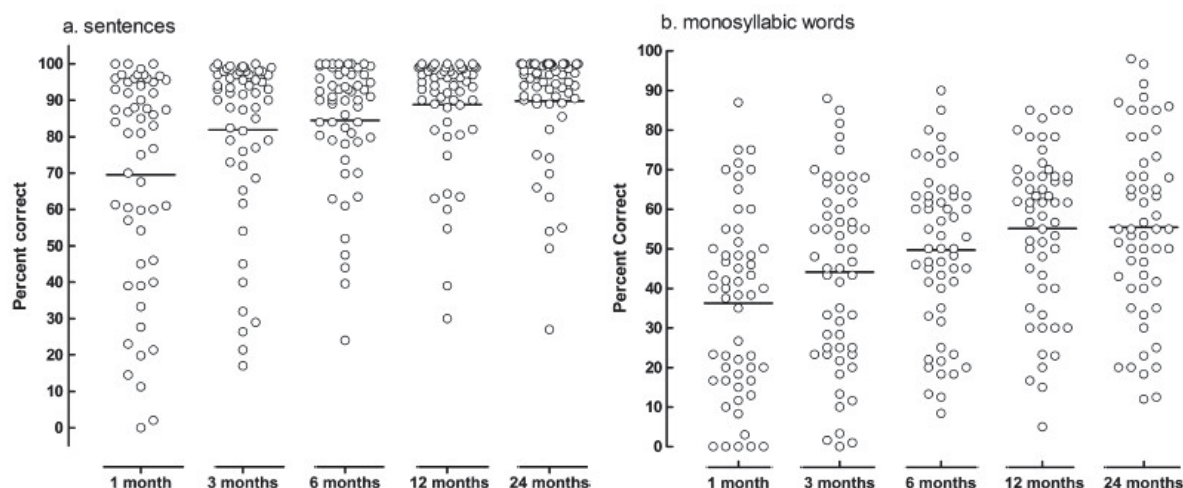


Figure 1. Pourcentage de reconnaissance de la parole (a : phrases, b : mots monosyllabiques) par 55 utilisateurs du système d'implant COMBI 40 avec la stratégie CIS. D'après Wilson et Dorman (2008).

La disparité des performances est tout aussi importante dans le bruit. Haumann et al. (2010) dans le centre de Hannover ont mesuré la résistivité au bruit chez 55 sujets qui ont ensuite été subdivisés en 5 groupes en fonction du processeur vocal qu'ils utilisent. Les auteurs utilisaient un test adaptatif. Ils justifient ce choix car, lorsqu'on sélectionne une population importante de patients, il est souvent difficile de réaliser des tests à RSB fixe. En effet, à RSB égal, les meilleurs patients vont plafonner à 100% d'intelligibilité alors que les plus mauvais seront proches de 0% de reconnaissance.

Dans cette étude, les stimuli étaient diffusés à différentes intensités, la parole était soit à $65 \text{ dB} \pm 10 \text{ dB}$, soit à $65 \text{ dB} \pm 15 \text{ dB}$, le bruit était un speech-shaped noise. Les auteurs recherchaient pour chaque sujet le seuil de reconnaissance de la parole (SRT = Speech Recognition Threshold), c'est-à-dire le RSB qui fournit une intelligibilité de 50% par rapport aux performances maximales dans le silence.

Leurs résultats montrent qu'il existe de très grandes différences de performances entre les meilleurs et les plus mauvais patients. Les patients les moins bons obtiennent des SRT de 17,3 dB tandis que les meilleurs ont des SRT de -4,8 dB, c'est-à-dire qu'ils résistent à des situations où le niveau de bruit est presque 2 fois plus fort que le niveau de parole (figure 2).

Ces résultats montrent une fois de plus une grande disparité entre les patients. Certains auront des performances dégradées pour un niveau de bruit relativement bas, de l'ordre de 8 fois plus faible que le niveau de parole, tandis que d'autres obtiennent des performances similaires alors que le bruit est presque 2 fois supérieur au niveau de parole.

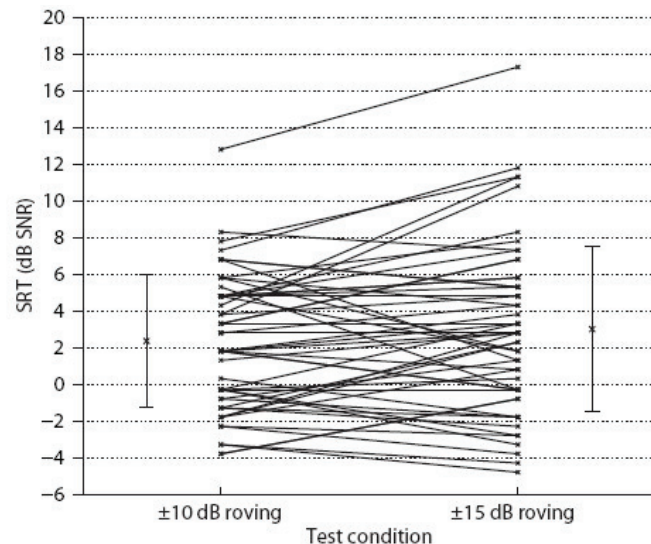


Figure 2. Seuil de reconnaissance de la parole (SRT) pour 55 sujets implantés cochléaires. Résultats individuels, moyens (à gauche et à droite) et déviations standard (barres d'erreurs) en fonction de la condition testée. D'après Haumann et al. (2010).

Les performances, peuvent-elles être corrélées à l'activité cérébrale des sujets ?

Observant une telle variabilité inter-individuelle pour les tests de compréhension de la parole, on peut se demander si cette variabilité peut être expliquée par des différences d'un point de vue du traitement de l'information. Pour répondre à cette question, divers travaux de recherche ont investigué les éventuelles relations entre les mesures perceptives et l'activité cérébrale chez les patients implantés. Généralement, ces travaux ont couplé des tests d'intelligibilité de la parole à des mesures électrophysiologiques ou à de l'imagerie.

En électrophysiologie, les premières études de Abbas et Brown (1991a, b) qui ont relié des mesures EABR aux performances de reconnaissance de la parole, n'ont pas permis de mettre en évidence une corrélation forte entre l'intelligibilité et les mesures dérivées des potentiels recueillis.

Plus tard, Groenen et al. en 1996 essayent de relier des mesures de potentiels évoqués auditifs au score de perception de la parole sur un pool de 7 patients divisés en deux groupes : 3 patients bons (compréhension entre 85 et 90%) et 4 patients considérés comme moyens (compréhension entre 67 et 76%). Leurs mesures montrent des latences P300 en moyenne 90 ms plus courtes chez le groupe de patients bons, par rapport au groupe de patients moyens. La plupart des amplitudes et des latences des pics N1, P2 et P300 des patients bons sont dans la normale. Tandis que chez tous les patients qui ont des scores moyens, l'amplitude du pic P2 et la latence P300 sont en dehors de la normale (figure 3).

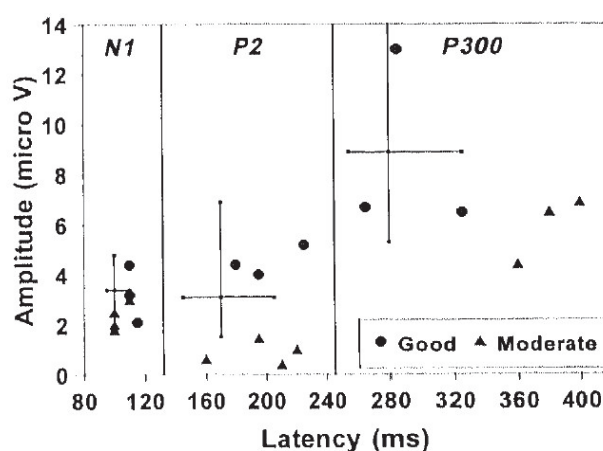


Figure 3. latences et amplitudes dérivées des mesures P300. Les valeurs de N1 et P2 étaient obtenues à partir des traces moyennes de stimuli tone bursts à 500 Hz. Les résultats indiqués par de cercles représentent les patients bons et les triangles représentent les patients moyens. Pour référence, la figure montre des résultats obtenus chez un groupe normo-entendant ($n = 11$), les lignes indiquent les gammes de valeurs moyennes obtenues pour les amplitudes et pour les latences. D'après Groenen et al. (1996).

En 1998, cette même équipe (Makhdoum et al., 1998) a mesuré les performances de perception des spondees (mots dissyllabiques) et les potentiels évoqués auditifs de latence précoce (EABR), moyenne (EMLR) et tardive (EALR), chez 15 patients implantés cochléaires utilisateurs du système Nucleus. Ils ont tenté de corrélérer les mesures perceptives aux mesures électrophysiologiques. Leur analyse montre que les scores de perception de la parole sont significativement corrélés à la latence de l'onde P2 ($r = -0,64$; $p < 0,01$; $n = 15$), mais ils ne sont corrélés avec la latence d'aucune autre pic. Les scores de perception de la

parole se révèlent également être corrélés à l'amplitude du complexe N1-P2 (figure 4 : $r = 0,53$; $p < 0,05$; $n = 15$).

N'observant pas de relation avec les ondes précoces et de latence moyenne, les auteurs en déduisent que les relais périphériques, n'auraient pas d'influence sur les capacités de perception de la parole avec l'IC.

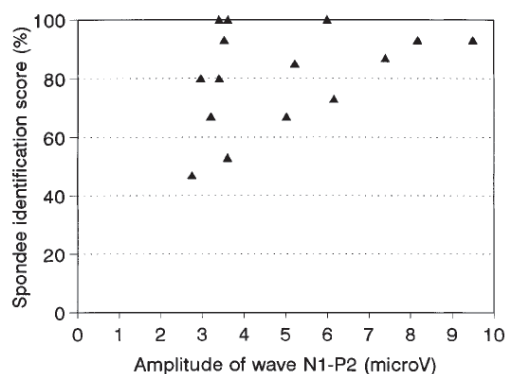


Figure 4. Relation entre les amplitudes des ondes N1-P2 et le score d'identification des spondees. D'après Makhdoum et al. (1998).

Ce résultat sera remis en question l'année suivante par Gallégo et al. (1998) qui trouvent des relations significatives entre les mesures EABR et la perception de la parole chez 17 sujets implantés avec le système Digisonic de MXM. Leurs données montrent que la reconnaissance des phonèmes est corrélée à la latence de l'onde V ($r = 0,57$; $p < 0,05$), à l'intervalle entre les ondes II-V ($r = 0,68$; $p < 0,005$) et à l'intervalle entre les ondes III-V ($r = 0,69$; $p < 0,005$). La latence de l'onde V et l'intervalle entre les ondes II-V apparaissent fortement corrélés avec l'intervalle entre les ondes III-V ($r = 0,77$ et $r = 0,90$, respectivement ; $p < 0,001$). Parmi toutes les variables de l'EABR prises en compte, leur analyse montre que l'intervalle entre les latences des ondes III-V (ce qui correspond au délai de propagation de l'influx nerveux entre le noyau cochléaire et le colliculus inférieur) est le plus puissant prédicteur des scores de reconnaissance de la parole (sans lecture labiale), pouvant compter jusqu'à 48% de la variance dans les données. Cette relation est représentée sur la figure 5.

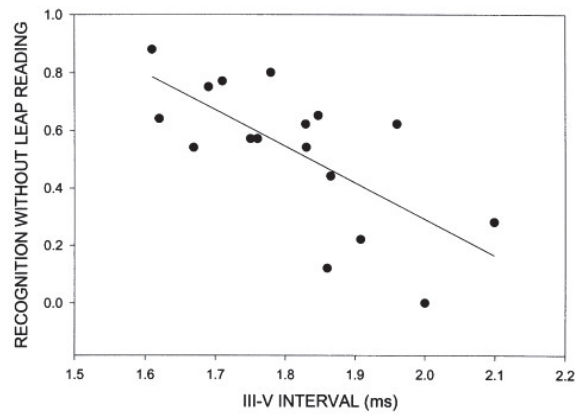


Figure 5. Relation entre l'intervalle de latence des ondes III-V de l'EABR et les scores de compréhension de la parole. D'après Gallégo et al. (1998).

Les auteurs supposent que la mauvaise compréhension de la parole serait liée à un allongement anormal du temps de propagation entre les structures nerveuses qui génèrent les ondes III et V. L'onde III est communément associée à l'activité du noyau cochléaire ipsilatéral, tandis que les générateurs du complexe IV-V seraient plus diffus et impliqueraient en particulier l'activité du colliculus inférieur controlatéral. On peut ainsi supposer que l'origine de cet allongement dans les temps de propagation résulte d'une dégénération neurale.

De plus, le fait que la compréhension soit liée à la latence de l'onde V et à l'intervalle entre les ondes III-V et ne soit pas liée à des ondes de latence plus précoce laisse supposer que les performances ne sont pas liées aux structures les plus périphériques du système auditif, mais bien aux premiers relais du traitement de l'information.

Quelques années plus tard, Maurer et al. (2002) ont utilisé un protocole sensiblement proche de celui de Gallégo et al. (1998) à la différence qu'ils ont couplé les résultats de leurs patients à des mesures de potentiels évoqués tardifs. Trente patients ont participé à cette étude : 15 bons sujets (intelligibilité des phonèmes entre 90% et 100%) et 15 moins bons (intelligibilité < 85%). En accord avec les résultats de Groenen et al. (1996) et de Makhdom et al. (1998), le seul facteur qui différencie les patients des deux groupes de patients est la latence de l'onde P2, l'onde se manifestant plus tôt (de 32 ms en moyenne) chez le groupe de patients qui ont les meilleures performances ($p = 0,016$). Les auteurs mettent en évidence une relation faible

mais significative entre cette variable électrophysiologique et le score d'intelligibilité de la parole ($r^2 = 17\%$; $p < 0,05$).

En ce qui concerne les mesures en imagerie, elles ont également permis de corrélérer les scores de perception de la parole aux mesures de l'activation cérébrale en réponse à des stimuli sonores langagiers.

Pour comprendre les mécanismes corticaux du traitement des signaux de parole codés par l'implant, Fujiki et al. (1999) ont évalué par des mesures TEP la corrélation entre l'activité cérébrale durant l'activation de parole avec les scores de perception de la parole chez 14 sourds post-linguaux implantés cochléaires. Les patients utilisaient l'implant Nucleus et la stratégie de codage SPEAK. Le flux sanguin local (rCBF = regional cerebral blood flow) était mesuré en absence de stimulation sonore et en réponse à des sons de parole. Les tests d'audiométrie vocale mesuraient la reconnaissance des voyelles, la reconnaissance des consonnes ainsi que le suivi de la parole (speech tracking). Les régions d'intérêt pour l'analyse étaient l'aire auditive primaire, l'aire auditive associative et l'aire de Broca.

Les analyses statistiques n'ont montré aucune corrélation entre les scores de perception de la parole et l'augmentation du rCBF dans les aires auditives primaires et dans l'aire de Broca, mais une corrélation positive est observée dans l'aire auditive associative. Dans l'aire auditive associative gauche (figure 6), le coefficient de corrélation avec la perception des voyelles est $r = 0,546$ ($p < 0,05$) et le coefficient pour le score de suivi de parole est $r = 0,657$ ($p < 0,05$). Enfin une forte corrélation est obtenue entre l'augmentation du rCBF et l'intelligibilité des consonnes ($r = 0,743$; $p < 0,01$). Dans l'aire auditive associative droite, une corrélation positive est observée, seulement entre l'augmentation de rCBF et l'intelligibilité des consonnes ($r = 0,576$, $p < 0,05$).

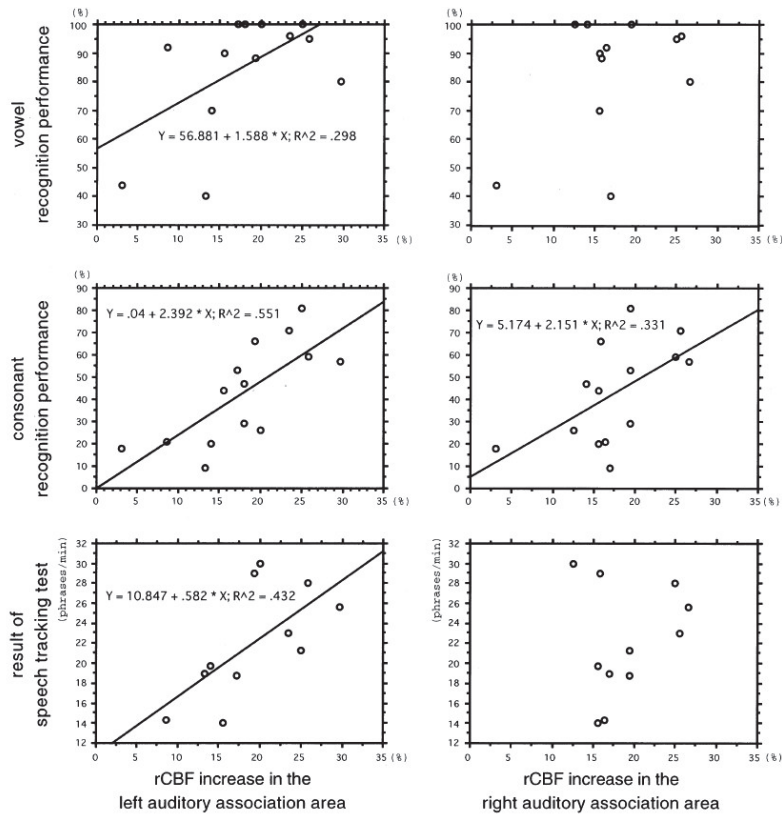


Figure 6. Corrélations entre l'augmentation du rCBF et les performances de perception de la parole pour 14 sujets implantés. A gauche : corrélations obtenues dans l'aire auditive associative gauche ; à droite : corrélations obtenues dans l'aire auditive associative droite. Dans l'aire associative droite, seule l'intelligibilité des consonnes est significativement corrélée à l'augmentation du rCBF. D'après Fujiki et al. (1999).

En 2001, Lee et al. ont sélectionné 10 sourds prélinguaux chez qui ils ont mesuré par TEP l'étendue des aires hypométaboliques, dans le cortex auditif primaire (BA 41) et dans le cortex auditif associatif (BA 42 et BA 22) des deux hémisphères, avant l'implantation cochléaire. Une fois implantés, les sujets étaient soumis à des tests d'audiométrie vocale à l'aide d'un équivalent coréen des phrases CID. Les patients implantés montraient une très large disparité de capacités auditives, certains stagnaient à 0% et d'autres plafonnant à 100% (moyenne = 42%). Les auteurs ont donc cherché à corréler les performances aux mesures TEP avant l'implantation et ils ont mis en évidence une corrélation positive ($r = 0,81$; $p < 0,005$; $n = 10$) entre le score de capacité auditive et la taille de l'aire hypométabolique dans BA 41, 42 et 22. Cette relation est représentée dans la figure 7.

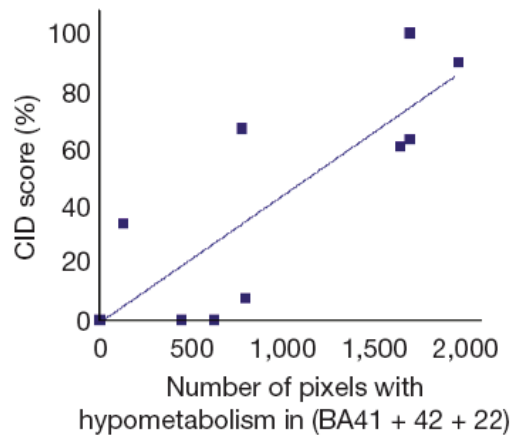


Figure 7. Résultats individuels des scores de perception de la parole (CID) en fonction de l'étendue de l'aire hypométabolique dans le cortex auditif primaire (BA41) et le cortex auditif associatif (BA 42, BA22) des deux hémisphères chez des patients sourds prélinguaux. L'étendue des aires hypométaboliques pré-implantation est significativement corrélée avec le score CID post-implantation. D'après Lee et al. (2001).

Quelques années plus tard, Green et al. (2005), avec la même technique d'imagerie, proposent eux aussi de corréliser l'activation cérébrale aux performances avec l'implant. Dix-huit patients implantés cochléaires étaient enrôlés dans leur étude : 6 implantés récents et 12 implantés expérimentés divisés en 2 sous-groupes, 4 ayant un score inférieur à 25% au test BKB et 8 ayant un score meilleur que 60%. L'intelligibilité de la parole pour la reconnaissance des phrases BKB variait de 0 à 100%.

Les auteurs notent que l'activité cérébrale était significativement plus importante chez le groupe de bons patients par rapport au groupe de patients qui a de mauvaises performances ($p = 0,034$ dans les aires auditives primaires ; $p = 0,0085$ dans les aires associatives ; $p = 0,0085$ dans toutes les aires auditives confondues).

Les analyses montrent une corrélation significative entre les performances de perception de parole et l'activité corticale dans le cortex auditif primaire ($r = 0,54$; $p < 0,05$) et dans le cortex associatif ($r = 0,67$; $p < 0,01$). Mais la corrélation est la plus forte entre la perception de la parole et l'activité corticale dans toutes les aires auditives ($r = 0,695$; $p < 0,005$). Lorsque le groupe de bons patients est analysé indépendamment, il y a une corrélation significative ($r = 0,77$; $p < 0,025$) entre la perception de la parole et l'activité corticale dans toutes les aires auditives. Ces relations sont représentées figures 8.

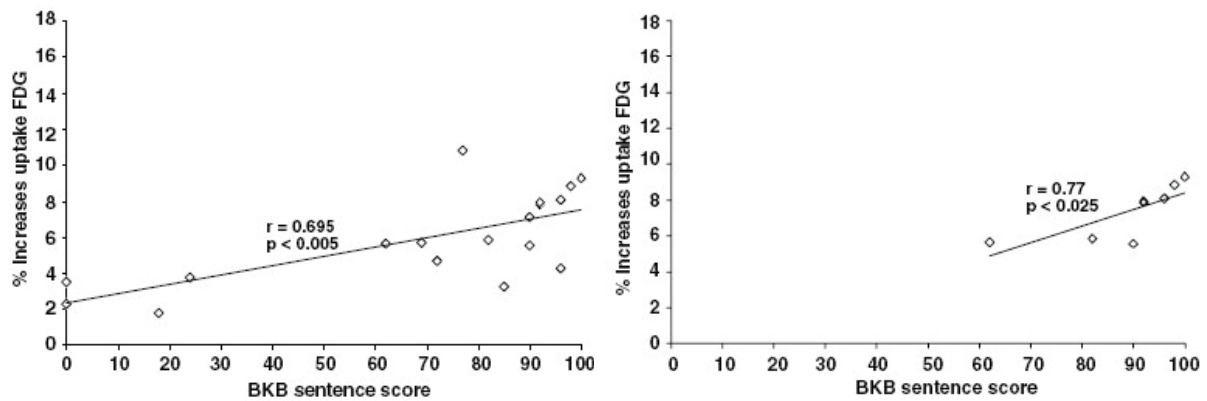


Figure 8. Corrélations entre la perception de la parole et l'activité corticale (aires auditives primaires et associatives), à gauche chez 18 adultes implantés cochléaires, à droite chez les 8/18 patients expérimentés ayant les meilleures performances. D'après Green et al. (2005).

D'après ces études, nous constatons que la dispersion observée ne se limite pas seulement aux scores de perceptions de la parole. Il existe bien des corrélats entre les mesures objectives et les performances. Quelles sont alors les causes de l'allongement des latences observé en mesures EABR et les causes de la diminution de l'activité cérébrale de l'aire auditive chez certains patients implantés?

Dans une seconde partie de leur étude, Green et al. (2005) cherchent à relier l'activité cérébrale des sujets à leur durée de surdité. Les auteurs affirment trouver une corrélation négative entre la durée de surdité et l'activation cérébrale, et entre la durée de surdité et les performances des patients (figure 9). Leurs patients avaient des durées de surdité allant de 1 à 51 ans. A partir des données individuelles de durées de surdités et du score BKB, nous avons pu retrouver la corrélation entre l'intelligibilité et la durée de surdité, ce qui nous conduit à des valeur de $r^2 = 81\%$ et $p < 0,001$.

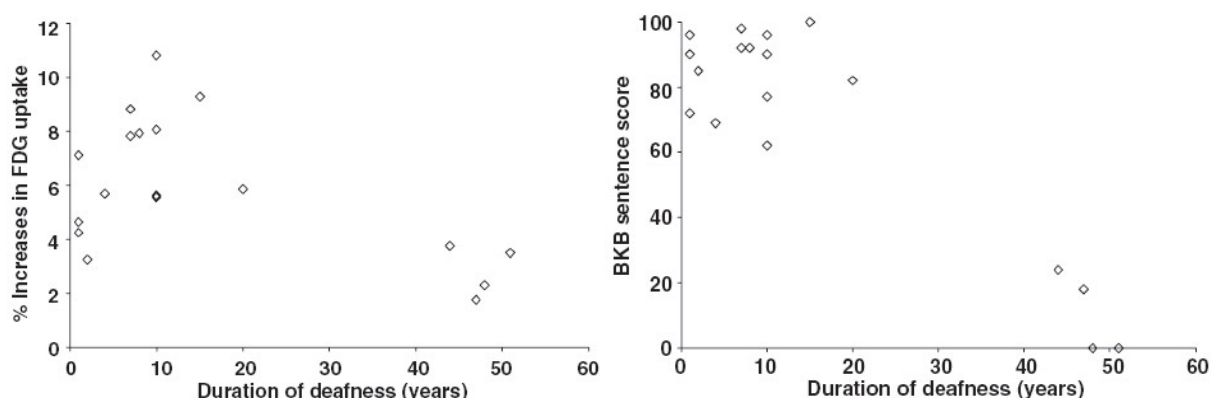


Figure 9. Activité cérébrale dans toutes les aires auditives (à gauche) et reconnaissance des phrases BKB (à droite), en fonction de la durée de surdité chez 18 implantés. D'après Green et al. (2005).

L'effet de la durée de surdité sur les mesures dérivées des potentiels évoqués a été clairement mis en évidence en 2007 grâce aux travaux de Guiraud et al. (2007a ; 2007b). Les auteurs ont mesuré les potentiels évoqués auditifs de latence précoce et tardive, respectivement chez 8 et 14 patients implantés cochléaires utilisateurs du système HiRes 90K de Advanced Bionics. Le but de cette étude était de rechercher les facteurs qui pouvaient être liés à l'allongement des latences des ondes, en particulier pour les ondes EABR IIIe et Ve, ce qui est observé par exemple dans Gallégo et al. (1998) et les ondes ELAR N1 et P2, ce qui est observé par exemple dans Maurer et al. (2002).

La première étude (Guiraud et al., 2007a) a mis en évidence un effet significatif de la durée de surdité sur la latence de l'onde IIIe ($p < 0,05$) et sur la latence de l'onde Ve ($p = 0,041$), mais l'effet n'est pas significatif sur l'intervalle IIIe-Ve ($p = ns$). La figure 10 illustre l'influence de la durée de surdité (inférieure ou supérieure à 7 ans) sur les latences des ondes IIIe (a), Ve (b) et l'intervalle IIIe-Ve (c).

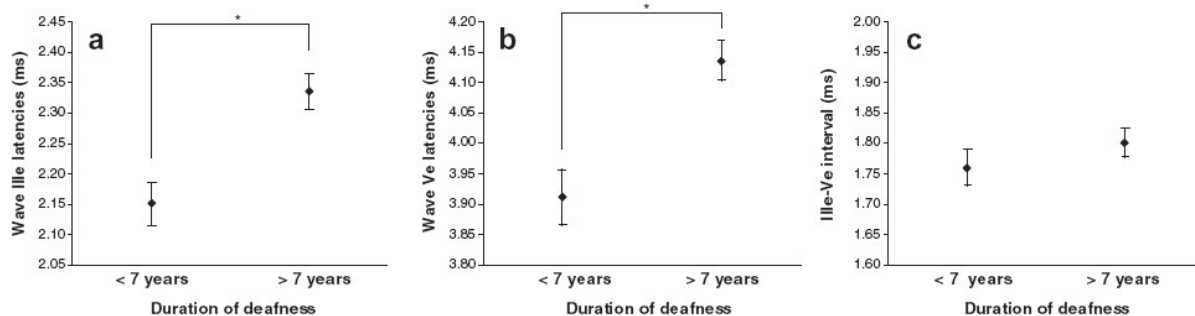


Figure 10. Latences moyennes des ondes EABR IIIe (a), Ve (b), et intervalle IIIe-Ve (c) moyennées à travers tous les sujets et tous les sites des électrodes, en fonction de la durée de surdit . D'apr s Guiraud et al. (2007a).

Cette  tude a  galement mis en  vidence un effet significatif des seuils audiom triques sur la latence de l'onde IIIe ($p = 0,013$) et sur la latence de l'onde Ve ($p = 0,004$), mais pas sur l'intervalle IIIe-Ve ($p = ns$). La figure 11 montre les latences des ondes IIIe (a), Ve (b) et l'intervalle IIIe-Ve (c) en fonction des seuils audiom triques (inf rieurs   100 dB HL pour le groupe 1 et sup rieurs   100 dB HL pour le groupe 2).

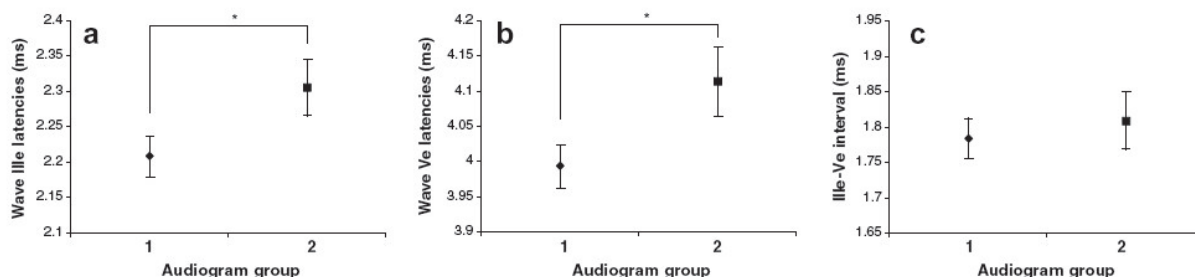


Figure 11. Latences moyennes des ondes EABR IIIe (a), Ve (b), et intervalle IIIe-Ve (c) moyenn es   travers tous les sujets et tous les sites des  lectrodes, en fonction du niveau d'audition avant l'implantation. Groupe 1 : perte auditive moyenne inf rieure   100 dB HL ; groupe 2 : perte auditive moyenne sup rieure   100 dB HL. D'apr s Guiraud et al. (2007a).

La seconde  tude (Guiraud et al., 2007b) a montr  que ces 2 effets avaient  galement une influence sur les latences des ondes tardives. L'analyse statistique montre un effet significatif des seuils audiom triques sur la latence de l'onde N1 ($p < 0,05$) et une tendance sur la latence de l'onde P2 ($p = 0,056$), mais pas sur l'intervalle N1-P2 ($p > 0,05$). La figure 12 montre les

latences des ondes N1 (a), P2 (b) et l'intervalle N1-P2 (c) en fonction des seuils audiométriques (inférieurs à 100 dB HL pour le groupe A1 et supérieurs à 100 dB HL pour le groupe A2).

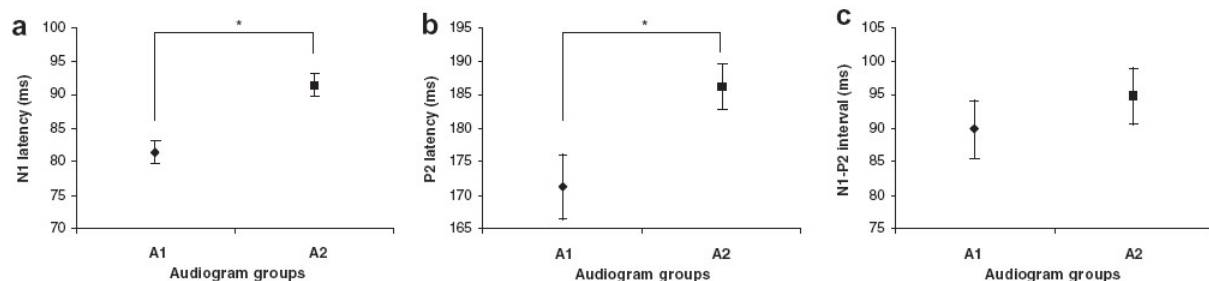


Figure 12. Latences moyennes des ondes ELAR N1 (a), P2 (b), et intervalle N1-P2 (c) moyennées à travers tous les sujets et tous les sites des électrodes, en fonction du niveau d'audition avant l'implantation. Groupe A1 : perte auditive moyenne inférieure à 100 dB HL ; groupe A2 : perte auditive moyenne supérieure à 100 dB HL. D'après Guiraud et al. (2007b).

Cette seconde étude a montré aussi un effet significatif de la durée de surdité sur la latence de l'onde N1 ($p < 0,05$), sur la latence de l'onde P2 ($p < 0,05$), et sur l'intervalle N1-P2 ($p < 0,05$). La figure 13 illustre l'influence de la durée de surdité (inférieure à 8 ans pour le groupe D1 ou supérieure à 8 ans pour le groupe D2) sur les latences des ondes N1 (a), P2 (b) et l'intervalle N1-P2 (c). Les auteurs constatent dans cet article que le groupe de sujets qui a la durée de surdité la plus courte est également celui qui a les meilleures performances en perception de la parole.

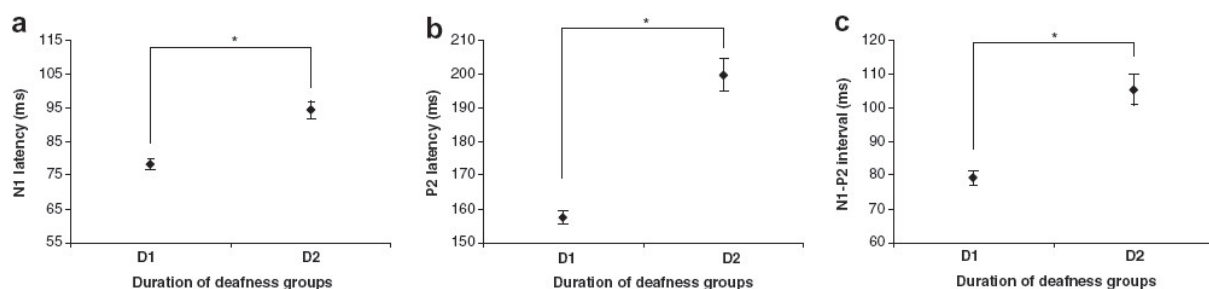


Figure 13. Latences moyennes des ondes ELAR N1 (a), P2 (b), et intervalle N1-P2 (c) moyennées à travers tous les sujets et tous les sites des électrodes, en fonction de la durée de surdité. Groupe D1 : les sujets ont moins de 8 ans de surdité et ont de meilleurs scores de perception de la parole ; groupe D2 : les sujets ont une durée de surdité supérieure à 8 ans et ont des scores de perception de la parole plus mauvais. D'après Guiraud et al. (2007b).

De même, dans l'article de Lee et al. (2001), les auteurs montrent que les sujets qui ont les plus mauvaises performances sont également ceux qui ont les surdités les plus anciennes. Ce constat laisse penser que suite à une longue période sans afférence sensorielle, l'aire auditive a probablement été colonisée par une autre modalité sensorielle et ne répondrait plus aux afférences auditives.

Dans la plupart des cas, la surdité touche d'abord les hautes fréquences et la majorité des patients adultes implantés ont eu une surdité évolutive. La privation sensorielle due à des lésions des récepteurs périphériques (ici les cellules basales de la cochlée) conduit généralement à une réorganisation massive des cartes sensorielles dans le cortex primaire. Une telle réorganisation résulte en une colonisation du territoire cortical déafférenté, qui répondait initialement à une afférence sensorielle donnée, par les aires adjacentes du cerveau impliquées dans le traitement d'entrées sensorielles différentes ou identiques. Dans la modalité auditive, suite à des lésions partielles sur la base de la cochlée, plusieurs études reportent des changements tonotopiques corticaux caractérisés par une sur-représentation des fréquences adjacentes à la fréquence de coupure de la zone lésée. Cette réorganisation post-lésion a été montrée dans un premier temps chez les animaux (Robertson et Irvine, 1989 chez le cochon d'inde par des lésions cochléaires mécaniques, Schwaber et al., 1993 chez le singe macaque et Kakigi et al., 2000 chez le chinchilla par l'administration de drogues ototoxiques) et plus tard chez l'humain (McDermott et al., 1998 et Thai-Van et al., 2002 et 2003 par des mesures psychoacoustiques ; Dietrich et al., 2001 et Di Nardo et al., 2006 avec des mesures objectives par imagerie). D'après Kral et Tillein (2006), les fréquences situées dans la région frontière entre le tissu cochléaire lésé et sain deviendraient représentées dans des régions corticales, précédemment utilisées pour des fréquences qui sont désormais dans la partie lésée de la cochlée. Ceci résulterait en une expansion de la région cochléaire fonctionnelle dans la région endommagée (figure 14) et ces mécanismes plastiques centraux se mettent en place au fil du temps. C'est pourquoi il existe une relation entre l'activité cérébrale et la durée de privation.

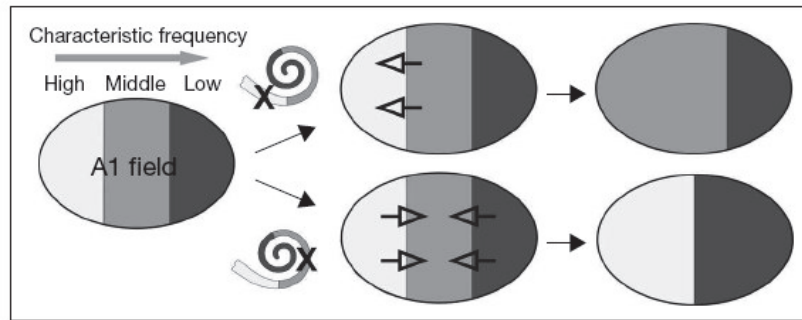


Figure 14. Effets d'un dommage cochléaire restreint sur les représentations corticales. Une lésion dans la région des hautes fréquences de la cochlée conduit, au niveau cortical, à une représentation étendue des moyennes fréquences. Une lésion dans la région des fréquences médiums de la cochlée conduit à une expansion des régions haute et basse fréquence au niveau du cortex auditif. D'après Kral et Tillein (2006).

La durée de surdité aurait donc des conséquences négatives sur l'influx nerveux dans le système auditif et sur l'activité des aires auditives. Quant à l'audition résiduelle, elle aurait une influence positive. D'après ce qui a été observé avec les mesures objectives, nous pouvons nous attendre à constater un effet similaire de la durée de surdité et de l'audition résiduelle pré-implantation sur les scores de reconnaissance de la parole. L'effet de ces deux facteurs est-il aussi important pour les tâches perceptives ? Et existe-t-il d'autres facteurs expliquant la disparité des performances ?

Disparité des performances en fonction de la durée de la surdité

Parmi les facteurs étudiés dans la littérature, nous trouvons généralement les caractéristiques propres aux patients, telles que l'âge d'apparition de la surdité (pré-, péri-, ou post-linguale), l'âge auquel l'implant a été mis en place, le degré d'audition résiduelle avant l'implantation, la durée d'utilisation de l'implant (l'expérience du patient), la durée de la surdité avant l'implantation ou encore la variable composite DURAGE introduite par Shipp et Nedzelski (1995), cette dernière correspondant à la proportion de vie sourde (= l'âge d'implantation ou la durée de la surdité divisé par l'âge du sujet). Enfin, nous trouvons des facteurs liés à la technologie de l'implant et à la stratégie utilisée pour coder le son.

En 1996, l'étude multicentrique de Blamey et al. a été une de première à intégrer un grand nombre de sujets. Les auteurs ont inclus 808 sujets (1064 mesures exploitables), ce qui leur a permis d'établir un modèle de performance auditive chez les adultes sourds post-linguaux implantés cochléaires, pour décrire l'effet de l'étiologie, de la durée de la surdité, de l'âge d'implantation, de l'âge d'apparition de la surdité et de la durée d'utilisation de l'implant. Ce modèle est présenté figure 15.

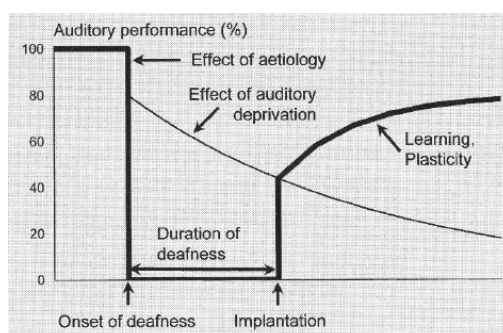


Figure 15. Modèle théorique en 3 étapes, de la performance auditive d'un adulte sourd post-lingual en fonction du temps. Les lignes en traits gras représentent les performances mesurables. Les lignes en traits fins représentent la performance auditive potentielle en fonction du temps avec un implant cochléaire, en fonction du temps écoulé après l'apparition de la surdité. D'après Blamey et al. (1996).

L'analyse des résultats montre un effet significatif de quatre facteurs principaux : la durée de la surdité, l'âge d'apparition de la surdité, l'étiologie et la durée de l'expérience avec l'implant. Ensemble, ces facteurs permettent d'expliquer 21% de la variance des données.

Le facteur qui compte pour la plus large proportion de la variance (13%) est la durée de la surdité (figure 17.a). La corrélation négative entre la durée de surdité et le pourcentage d'intelligibilité est en effet fortement significative ($r = -0,312$; $p < 0,0001$). Après avoir soustrait l'effet de la durée de privation, l'analyse des résidus révèle que l'âge d'implantation n'a qu'un léger effet négatif sur les performances, tandis que l'âge d'apparition de la surdité ne semble pas avoir d'impact sur les performances, pour les surdités apparues avant 60 ans.

Dix ans plus tard, dans une revue de littérature Cooper (2006) présente les résultats obtenus dans l'étude 'POCIA' (Predicting Outcomes of Cochlear Implantation in Adults) qui se réfère aux trois articles du UK Cochlear Implant Study Group (2004a, b, c). Cette étude multicentrique a été réalisée chez 317 patients à travers 13 centres d'implantation cochléaire du Royaume Uni, à l'aide de tests de performances et de questionnaires. Les facteurs pré-opératoires évalués étaient physiologiques (par exemple la durée de surdité ou l'audition résiduelle) et cognitifs (par exemple capacités de lecture labiale), mais les auteurs évaluaient également la motivation du sujet.

Comme dans les études citées précédemment, la variabilité des performances inter-patients est relativement grande (figure 17.f). Des analyses de régression entre les variables testées et les mesures de performances post-implantation ont montré des effets significatifs des différents facteurs, toutefois les coefficients de corrélation étaient généralement très faibles. Il ressort néanmoins de ces analyses que la plus forte corrélation avec la discrimination des phrases à 9 mois post-implantation est obtenue avec la variable « durée de surdité dans l'oreille implantée » ($r = -0,429$; $p < 0,001$) qui explique 18% de la variance, suivie de près par la durée de surdité dans l'autre oreille ($r = -0,289$; $p < 0,001$) qui explique 8% de la variance.

A Baltimore aux Etats-Unis, Leung et al. (2005) ont réalisé une étude prospective sur 749 sujets. Ils ont cherché à déterminer, à partir des données préopératives, les facteurs prédictifs des performances post-implantation. Parmi les facteurs étudiés, les auteurs ont pris en compte la durée de la surdité, le score d'intelligibilité des phrases, l'âge d'implantation et la proportion de vie sourde, c'est-à-dire la variable DURAGE décrite précédemment.

En accord avec ce qui a été décrit dans la littérature, les auteurs observent une très grande disparité des résultats, entre 0 et 100%. Les données individuelles de 627 patients sont présentées sur la figure 16. Cette figure montre le score post-implantation d'intelligibilité des mots CNC en fonction de l'âge d'implantation des sujets. Du fait de la large variabilité des résultats quelque soit l'âge des patients, les analyses statistiques n'ont pas mis en évidence une corrélation significative entre l'âge d'implantation et les performances post-implantation chez cette cohorte.

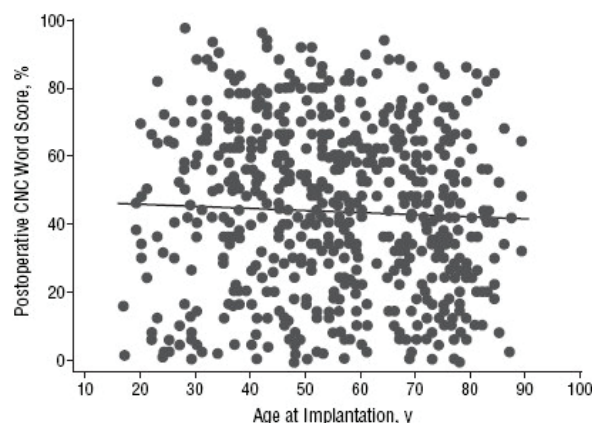


Figure X. Pourcentage d'intelligibilité des mots CNC post-implantation en fonction de l'âge d'implantation. Meilleur score obtenu durant la première année d'utilisation de l'implant chez 627 sujets. D'après Leung et al. (2005).

En ce qui concerne la corrélation des performances avec la durée de la surdité, l'analyse sur 523 sujets, tous âges confondus, montre une diminution générale de l'intelligibilité des mots lorsque la durée de la surdité augmente. Les 523 sujets ont été divisés en 2 groupes, en fonction de leur âge, supérieur ou inférieur à 65 ans. Une analyse chez les sujets de moins de 65 ans révèle que la diminution de la compréhension des mots est associée à l'augmentation de la durée de surdité ($n = 315$ sujets ; $p < 0,001$). Chez le groupe de plus de 65 ans, l'effet de la durée de la surdité n'est pas significatif ($n = 208$ sujets ; $p = 0,13$).

Enfin, la corrélation des performances avec le DURAGE a été mesurée avec 521 sujets (206 ont plus de 65 ans et 315 ont plus de 65 ans). Lorsque ce ratio augmente (lorsque le sujet a passé une grande proportion de sa vie sans entendre), le score post-implantation diminue. L'effet est particulièrement fort chez le groupe plus jeune avec une valeur de $p < 0,001$ (versus $p = 0,10$ dans le groupe plus vieux). D'après les auteurs, le ratio autour de 2/3 de vie sourde pré-implantation semble être une limite au-delà de laquelle, les patients implantés auront du mal à comprendre. Kessler et al. (1995) ont par exemple évalué l'intelligibilité avec l'implant Clarion chez 64 patients. Ils trouvent en particulier que des proportions de vie sourde supérieures à 60% sont associées à de plus mauvaises performances. De même, Geier et al. (1999) constatent que les patients qui ont été sourds pendant 60% de leur vie ou plus nécessitent beaucoup plus de temps pour atteindre leurs performances optimales avec l'implant, par rapport à des patients qui ont eu une proportion de vie sourde moins importante.

Récemment, Green et al. (2007) ont réalisé une étude rétrospective sur les résultats des patients implantés pendant les 14 premières années d'activités du centre d'implantation de Manchester (de 1988 à 2002) et 117 patients ont été sélectionnés pour leur analyse. Les auteurs proposaient d'évaluer si les performances post-implantation étaient liées aux facteurs sexe, âge d'implantation, durée de perte auditive, étiologie de la surdité, audition résiduelle avant l'implantation, type d'implant, stratégie de codage et nombre d'électrodes actives (exprimée en pourcentage d'électrodes insérées dans la cochlée).

Les résultats individuels représentés figure 17,g. montrent que la durée de la surdité avant l'implantation s'avère être le seul facteur prédictif des performances post-implantation, indépendamment des autres variables ($r = -0,557$; $p < 0,0001$) et ce facteur expliquerait 9% de la variabilité des résultats.

Leur analyse montre que l'âge d'implantation n'aurait pas d'effet sur les performances avec l'implant (en accord avec les résultats de Blamey et al., 1996 et de Leung et al., 2005 décrits précédemment). Ils trouvent en revanche que la durée de surdité est significativement corrélée avec les performances en condition audition seule ($r = -0,47$; $p < 0,01$), la durée de surdité permettrait donc d'expliquer 22% de la variance. Cette corrélation est d'autant plus importante lorsque la durée de surdité est exprimée en terme de proportion de vie sourde (DURAGE) et dans ce cas, 27% de la variance est expliquée par le modèle de régression proposé ($r = -0,52$; $p < 0,01$).

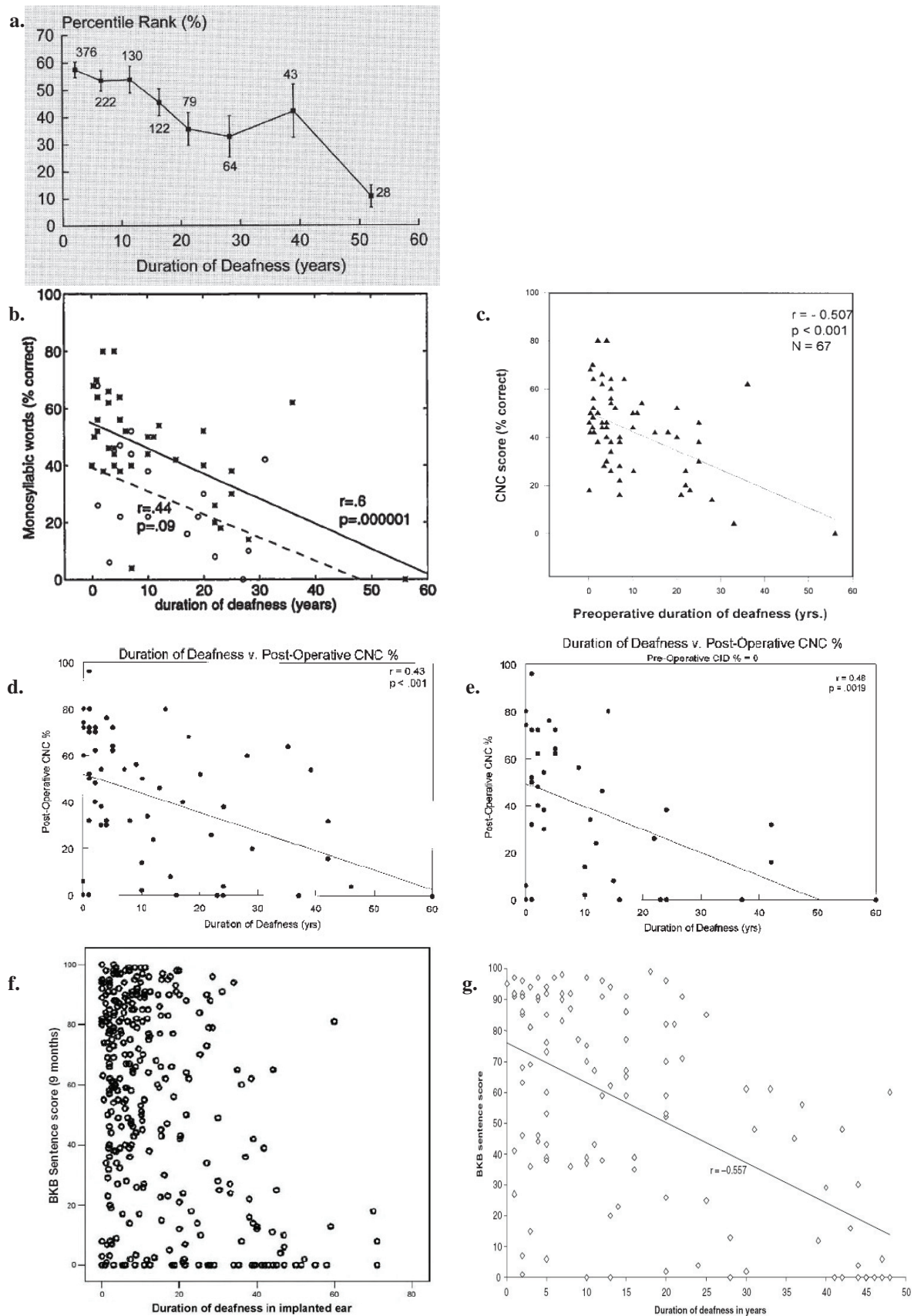


Figure 17. Intelligibilité avec l'implant cochléaire en fonction de la durée de surdité : a. d'après Blamey et al. (1996) chez 808 sujets (1064 mesures) ; b. d'après Rubinstein et al. (1999) chez 37

sujets (décrit ci-après) ; c. d'après Gomaa et al. (2003) chez 60 sujets (décrit ci-après) ; d. et e. d'après Friedland et al. (2003) respectivement chez 58 patients et chez 39/58 patients ayant une intelligibilité nulle avant l'implantation (décrit ci-après) ; f. d'après Cooper et al. (2006) chez 317 sujets ; g. d'après Green et al. (2007) chez 117 sujets.

D'après les différents auteurs, un des plus importants facteurs prédictifs des performances post-implantation seraient donc la durée de surdité profonde ainsi que la proportion de vie sourde. L'âge d'implantation, quant à lui, ne semble pas avoir d'influence majeure sur les résultats avec l'IC. Pour autant, la durée de privation qui apparaît comme étant le plus fort facteur prédictif n'explique pourtant que 9 à 37% de la variance des résultats selon les études. Ceci signifie que la variabilité est liée entre 60 et 90% à d'autres facteurs.

Disparité des performances en fonction de l'expérience du sujet, de la génération de l'implant, de la stratégie et des indications.

Généralement, les performances des patients implantés évoluent durant les 2 premières années d'utilisation. Selon Tyler et al. (1997), la compréhension avec l'IC augmente drastiquement durant les 9 premiers mois, puis il y a une évolution moins importante entre 19 et 30 mois de port. Certains sujets présentent une augmentation de leurs performances jusqu'à 4 ans d'utilisation, mais généralement une stabilité est atteinte au-delà de 30 mois. Ce constat est toujours d'actualité (cf Wilson et Dorman, 2008 : figure 1).

Adunka et al. (2008), dans une étude récente, aboutissent au même constat. Les auteurs ont étudié l'effet de l'audition résiduelle préopératoire sur la perception de la parole après l'implantation cochléaire. Ils ont comparé deux groupes de patients, un groupe de 29 patients présentant une audition résiduelle importante avant l'implantation et un autre groupe contrôle, constitué de 21 patients qui présentaient les mêmes caractéristiques que le premier groupe en terme d'âge et de durée de surdité, mais qui n'avaient pas d'audition résiduelle. Les sujets ont été soumis à des tests de compréhension des phrases (CUNY), des mots CNC, et des test HINT dans le silence et dans le bruit (RSB = +10 dB). Les scores d'intelligibilité sont mesurés entre 3 et 4 mois avant la chirurgie et post-implantation, au bout d'un mois, de 3 mois, de 6 mois et de 1 an. La figure 18 représente les résultats individuels des sujets, nous

constatons une évolution des performances tout au long de la première année d'utilisation de l'implant chez les deux groupes de sujets (avec et sans audition résiduelle).

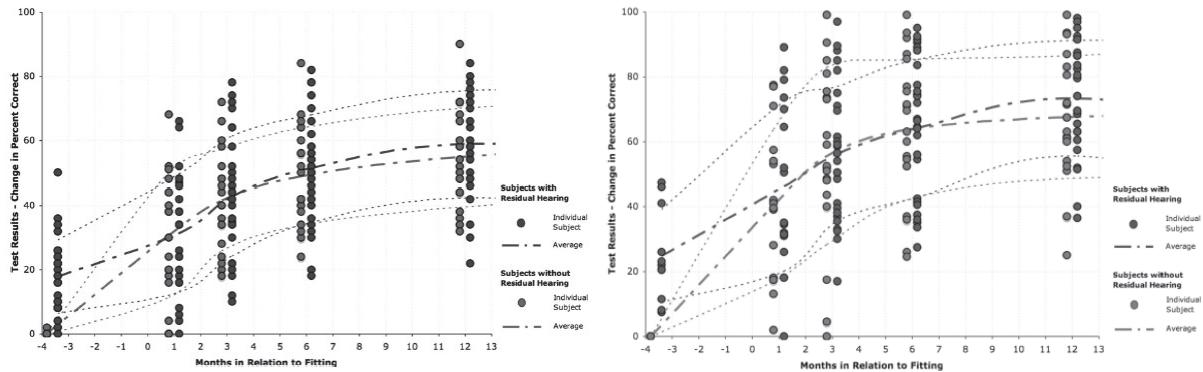


Figure 18. Scores individuels d'intelligibilité de la parole en fonction de la session de test entre 4 mois pré-implantation et 1 an post-implantation, pour les mots CNC à gauche et pour le test HINT à un RSB = +10 dB à droite. Les cercles noirs représentent les sujets ayant de l'audition résiduelle avant l'implantation, les cercles gris représentent les patients n'ayant pas d'audition résiduelle. Les résultats ne montrent pas de différence significative de performance entre les deux groupes pour les différents tests et les différentes sessions ($p > 0,05$). D'après Adunka et al. (2008).

En 2008, Krueger et al. ont comparé les performances de 864 patients implantés à Hannovre entre 1984 et 2008. Leurs sujets ont été classés en 5 groupes, en fonction de leur implant et de la stratégie de codage qu'ils utilisent. Le groupe 1 est composé de 107 patients implantés entre 1984 et 1990 qui utilisent principalement l'implant Nucleus 22 avec une stratégie d'ancienne génération type F0/F2. Le groupe 2 est composé de 94 patients implantés principalement avec le Nucleus 22 et la stratégie Multipeak. Le groupe 3 contient 104 sujets implantés avec le Nucleus 22 et Nucleus 24 avec la stratégie SPEAK (Spectral Peak), ainsi que des sujets implantés avec le Clarion 1.0 ou 1.2 avec une électrode radiale bipolaire utilisant la stratégie CIS ou SAS. Le groupe 4 est composé de 249 sujets implantés soit avec le Clarion 1.2 en bipolaire amélioré et la stratégie CIS ou SAS, soit avec le Nucleus 24 réglé en ACE. Enfin, le groupe 5 contient 310 patients utilisateurs de la technologie la plus récente, à savoir le CII ou le HiRes 90K de Advanced Bionics, le Freedom de Cochlear, et les C40+ et Pulsar de Med-El.

Les sujets ont été soumis à différents tests de compréhension de la parole, dans le silence et dans le bruit. Les résultats montrent que le groupe 1, qui a les implants les plus vieux, a des performances beaucoup moins bonnes que les autres groupes. Pour les sujets du groupe 1, les résultats augmentent de façon assez monotone au fil du temps, tandis que pour les patients des autres groupes qui ont des implants plus récents, les performances augmentent assez rapidement les premières années et atteignent un niveau stable au bout des 2 premières années d'utilisation de l'implant. Pour les tests de compréhension dans le bruit, seuls les groupes avec des implants récents (groupes 4 et 5) parviennent à obtenir des scores raisonnables. Les patients qui portent des implants plus anciens ont beaucoup plus de difficultés à comprendre dans le bruit (figure 19).

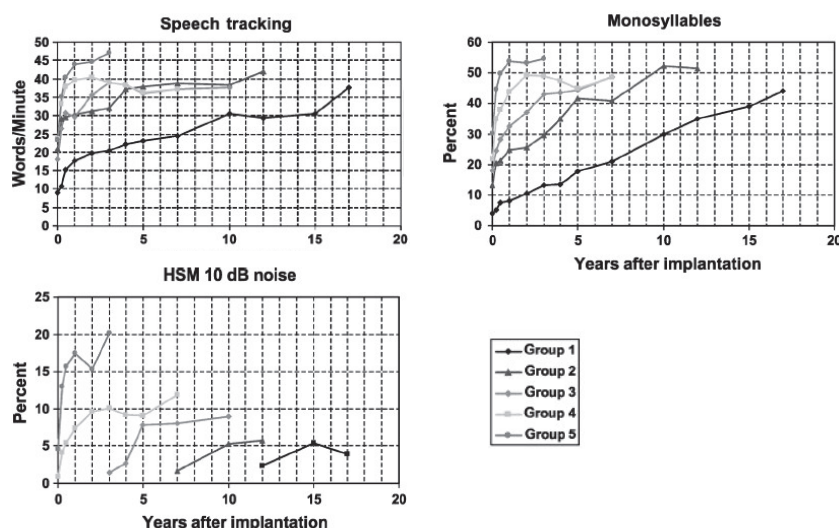


Figure 19. Haut gauche : score de suivi de la parole (en nombre de mots/minute) en fonction de la durée de port de l'implant. Haut droite : pourcentage d'intelligibilité des syllabes en fonction de la durée de port de l'implant. Bas gauche : Pourcentage d'intelligibilité des phrases dans le bruit, à un RSB de 10 dB, pour les 5 groupes de patients. D'après Krueger et al. (2008).

Quelques années auparavant, Rubinstein et al. (1999) avaient montré des résultats similaires (figure 20.a), suggérant que les processeurs les plus récents fournissaient de meilleurs résultats. Cependant, comme le soulignent les auteurs (Rubinstein et al., 1999 ; Krueger et al., 2008), depuis le début de l'implant cochléaire il y a eu non seulement une évolution des technologies (figure 20) mais également et surtout une évolution des indications. Les patients implantés le plus récemment ont également une meilleure audition résiduelle avant

l'implantation. Ce qui signifie que les meilleures performances observées chez les patients qui ont les implants les plus récents n'est pas purement attribuable à la stratégie de codage. Toutefois il n'est pas possible de dissocier ce qui est imputable à la technologie et ce qui est imputable à l'audition résiduelle.

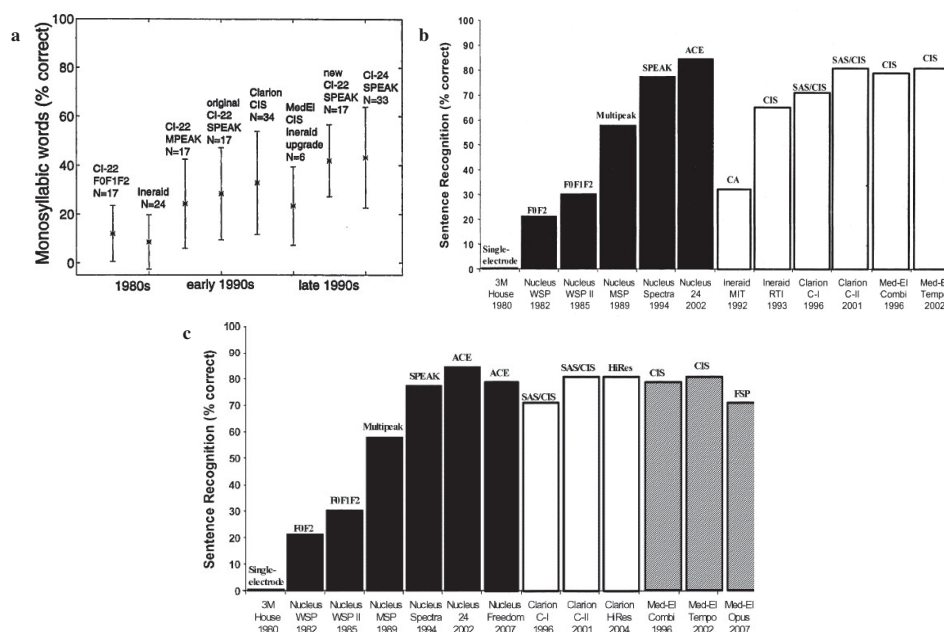


Figure 20. Performances moyennes des patients implantés en fonction de la stratégie de codage : a. d'après Rubinstein et al. (1999) pour les patients du centre d'Iowa ; b. et c. d'après Zeng (2004) et Zeng et al. (2008) avec les données publiées dans la littérature.

Au début de l'implant cochléaire, seuls les patients qui n'avaient aucune perception auditive pouvaient être candidats à l'implantation. Il y a une vingtaine d'année encore, il n'était pas rare de trouver des sourds congénitaux implantés à l'âge adulte (Busby et al., 1992). Depuis le milieu des années 80, les indications se sont relaxées et les centres d'implantation ont tendance à implanter des malentendants qui ont de plus en plus d'audition résiduelle dans les basses fréquences. Nous pouvons le constater sur la figure 21 qui montre les seuils préopératoires à 250 Hz de tous les patients qui ont été implantés à l'Université d'Iowa entre 1986 et 1999 (Tyler et al., 2002).

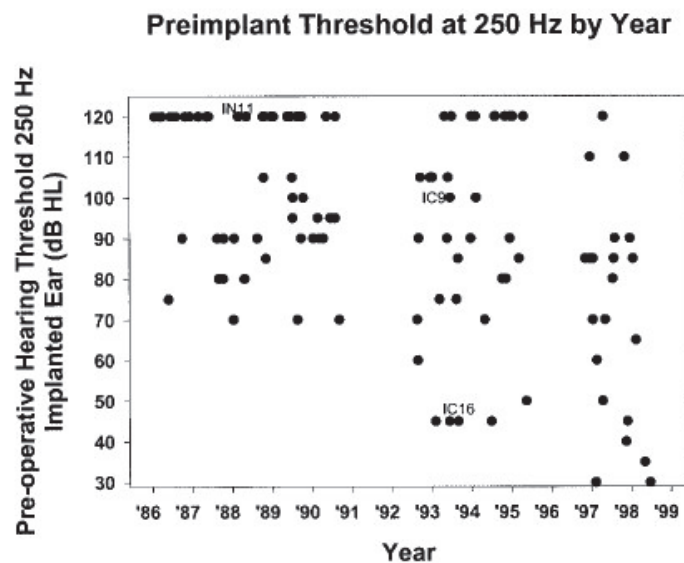


Figure 21. Seuil d'audition pré-implantation à 250 Hz dans l'oreille implantée, pour tous les patients implantés cochléaires adultes de 1986 à 1999 à l'Université d'Iowa. D'après Tyler et al. (2002).

Aujourd'hui, l'implant cochléaire peut être proposé aux personnes présentant une perte auditive sévère, c'est-à-dire des personnes qui ont des reliquats d'audition mais également de compréhension. Actuellement en France, d'après la Haute Autorité de Santé, lorsque l'intelligibilité en mots dissyllabiques devient inférieure à 50%, à 65 dB HL avec des aides auditives numériques bien réglées, nous pouvons considérer que l'implant cochléaire fournira un bénéfice supérieur à l'aide auditive (Truy et al., 2009 ; Eshraghi et al., 2009).

D'après Rubinstein et al. (1999), Gomaa et al. (2003) et Friedland et al. (2003), la reconnaissance de parole résiduelle pré-implantation agirait comme une « facteur trophique » qui protégerait le ganglion spiral et/ou les voies auditives centrales de la dégénérescence. En d'autres termes, elle améliorerait les scores post-implant attendus. Cette hypothèse pourrait expliquer l'effet positif de l'audition résiduelle sur les réponses électrophysiologiques dans les travaux de Guiraud et al. (2007a,b).

Plusieurs études ont évalué l'influence de l'audition et de la compréhension pré-implantation sur les performances post-implantation. Par exemple van Dijk et al., en 1999 ont recherché les facteurs qui pourraient prédire les performances des patients post-implantation. Leur étude a été réalisée chez 37 sujets sourds post-linguaux. Avant l'implantation, les seuils

audiométriques en conduction aérienne étaient mesurables chez 25 patients pour le 1 kHz et chez 29 patients pour le 250 Hz. Les auteurs ne trouvent pas de corrélation significative entre le seuil de perception à 250 Hz et le score d'intelligibilité post-implantation. Ils supposent que ce résultat peut être expliqué par le fait qu'à 250 Hz, à forte intensité, les sujets répondent à une vibration plutôt qu'à un son, ce qui biaiserait les données. Le seuil à 1000 Hz quant à lui, apparaît significativement corrélé aux performances ($r = -0,42$; $p < 0,05$), et permet d'expliquer 18% de la variance.

Les analyses montrent par ailleurs que les résultats des patients sont liés à leur score de reconnaissance pré-implantation. Ils sont liés au score de reconnaissance des sons environnementaux ($r = 0,38$; $p < 0,05$), ainsi qu'au score d'intelligibilité des mots dissyllabiques (spondees : $r = 0,39$; $p < 0,05$). Ceci signifie que les patients qui comprennent le mieux avant l'implantation sont susceptibles d'être les patients qui auront la meilleure intelligibilité une fois implantés.

La même année, l'étude d'Iowa a présenté des résultats similaires (Rubinstein et al., 1999). Cette étude rétrospective conduite sur 46 sujets utilisait des modélisations statistiques pour lier les données préopératoires aux scores de reconnaissance avec l'implant. Les tests utilisés en préopératoire mesuraient l'intelligibilité des phrases (CID sentences), tandis que les tests post-implantation mesuraient l'intelligibilité des mots monosyllabiques (CNC words), ceci afin d'éviter d'obtenir des scores à 0% et à 100%. Dans leur article, ils présentent les résultats obtenus chez 46 patients utilisateurs de l'implant Cochlear CI-22 et CI-24 avec la stratégie SPEAK. Les sujets sont séparés en deux groupes, la différence majeure entre les deux groupes étant l'audition résiduelle pré-implantation (meilleure chez le groupe le plus récent). Pour le groupe initial de sujets utilisateurs du CI-22, l'intelligibilité des mots CNC est négativement corrélée avec la durée de surdité ($r = 0,44$; $p = 0,09$). Pour l'autre groupe incluant les nouveaux patients CI-22 et les patients CI-24, les performances sont également corrélées avec la durée de perte avec un coefficient de corrélation plus important ($r = 0,6$; $p < 0,001$). Dans ce cas, la durée de surdité expliquerait 36% de la variance des résultats. Pour les deux groupes de patients, la pente de la droite de régression est identique : chaque année additionnelle de surdité profonde conduit à une diminution des performances de 0,9% pour la compréhension des mots CNC (figure 17.b. : les astérisques incluent les nouveaux patients CI-22 et CI-24 correspondant à la droite de régression en trait plein. Les cercles incluent les patients initiaux

CI-22, la droite de régression est tracée en traits pointillés). Sur les résultats présentés, la différence de performance entre les deux groupes est de 14%. La différence de critère d'implantation étant l'audition résiduelle, on peut supposer que l'audition résiduelle a une influence sur les résultats post-implantation.

Dans une deuxième partie de leur étude, Rubinstein et al. ont alors analysé l'effet de l'audition résiduelle sur les performances avec l'implant. Ils trouvent en effet une corrélation faible mais significative entre l'intelligibilité des phrases pré-implantation et l'intelligibilité des mots post-implantation (figure 23.a : $r = 0,34$; $p = 0,019$). La pente de la droite de régression indique que pour 2% additionnels de compréhension des phrases avant la chirurgie, une augmentation de 1% de la compréhension des mots CNC est attendue après l'implantation. Ainsi, pour une durée de surdité donnée, si un patient comprend 40% des phrases en pré-opératoire, il aura un score de reconnaissance des mots CNC supérieure de 20%, par rapport à un patient qui a la même durée de surdité, mais qui n'a aucune intelligibilité avant l'implantation.

A partir de ces observations, les auteurs ont construit un modèle bivarié pour modéliser les performances avec l'implant. La figure 22 montre l'effet de la surdité sur la compréhension moyenne avec l'implant pour deux patients hypothétiques : un qui n'a aucune compréhension avant la chirurgie et un qui comprend 40% des phrases, la maximum approuvé à l'époque par la FDA dans les critères d'implantation. Le modèle montre une différence importante dans les pentes des deux droites ; le patient qui n'a aucune reconnaissance de la parole pré-implantation voit ses performances diminuer de 1,1% par année de surdité. Le patient avec 40% de compréhension avant la chirurgie voit ses performances diminuer de moitié, à une vitesse de 0,52% par an. Ceci met en évidence que l'audition résiduelle, mesurée en terme de reconnaissance des phrases a une importance capitale pour l'amélioration des effets délétères de la durée de surdité.

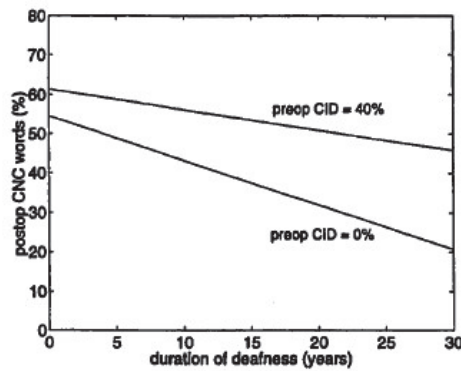


Figure 22. Prédiction du pourcentage de reconnaissance des mots post-implantation en fonction de la durée de surdit . Calculs pour deux patients hypoth tiques, un qui n'a aucune compr hension avant l'implantation et un qui a un score de compr hension de 40% avant l'implantation. D'apr s Rubinstein et al. (1999).

En 2003, Gomaa et al. ont men  une  tude chez 60 patients (dont 7 implant s bilat raux) avec une surdit  post-linguale de degr  s v re   profonde. Pour les m mes raisons que Rubinstein et al. (1999), les tests utilis s en pr op ratoire mesuraient l'intelligibilit  des phrases (CID sentences) et les tests post-implantation mesuraient l'intelligibilit  des mots monosyllabiques (CNC words). Leurs donn es pr sentent une large disparit  des performances avec l'implant qui est li e   la dur e de privation (figure 17.c). La relation est statistiquement significative ($r = 0,507$; $p < 0,001$). La pente de la droite de r gression ($-0,79$) indique que chaque ann e additionnelle de surdit  profonde conduit en moyenne   une diminution de 0,8% pour la compr hension des mots CNC post-implantation. Un de leurs sujets pr sentait toutefois des r sultats aberrants et a du  tre exclu de l'analyse. Leur nouvelle interpr tation conduit   une valeur de $r = 0,416$ (17% de la variance expliqu e) au lieu de 0,507 mais la relation est toujours significative ($p < 0,001$). La nouvelle droite de r gression a une pente de 0,73 sans ce sujet, ce qui signifie que chaque ann e additionnelle de surdit  est responsable d'une diminution de 0,7% de la perception de mots en post-op ratoire.

La figure 23.b montre la relation trouv e par Gomaa et al. entre l'intelligibilit  des phrases pr -implantation et l'intelligibilit  des mots post-implantation. Les auteurs trouvent une corr lation faible mais significative ($r = 0,267$; $p = 0,029$) entre les deux param tres. La droite de r gression a une pente de 0,44 qui indique que chaque augmentation de 10% pr -implantation conduit   une augmentation d'environ 4,4% du score post-implantation. En d'autres mots, un patient avec une intelligibilit  de 40% avant l'implantation aura

probablement un avantage de 17,6% post-implantation, par rapport à un patient qui n'avait pas de compréhension résiduelle.

Parallèlement aux travaux de Gomaa et al., Friedland et al. (2003) ont mené une étude similaire. Leur population était constituée de 58 patients sourds post-linguaux ayant une intelligibilité des phrases de 40% ou moins avant leur implantation. Les patients étaient majoritairement appareillés avec l'implant Clarion ou Nucleus. La figure 17.d montre le score obtenu pour la reconnaissance des mots CNC post-implantation en fonction de la durée de surdité. La corrélation est significative ($r = 0,43$; $p < 0,001$). Ce résultat est similaire aux résultats présentés par Rubinstein et al. (1999). Dans leur cohorte, Friedland et al. avaient 2/3 de leur patients (39/58) qui n'avaient aucune intelligibilité pré-implantation, il ont donc fait une analyse complémentaire sur ce sous-groupe de patients (figure 17.e) et le coefficient de corrélation est là encore significatif ($r = 0,48$; $p = 0,0019$). La pente de leur droite de régression montre une diminution des performances de 1% par année de surdité additionnelle, en accord avec les études de Rubinstein et de Gomaa présentées précédemment. Une analyse complémentaire montre que la corrélation est meilleure pour les patients appareillés avec le système Nucleus chez qui 21% de la variance du score CNC est expliquée par la durée de surdité, alors que pour les patients utilisateurs du système Clarion, seule 14% de la variance est expliquée.

Friedland et al. ont ensuite corrélé les performances pré- et post-implantation mais leur analyse n'a pas mis en évidence de relation significative entre les deux paramètres ($r = 0,20$; $p = 0,13$). Les auteurs ont alors supprimé les « star patients » de leur analyse et il apparaît une corrélation significative entre les score d'intelligibilité des phrases CID pré-implantation et l'intelligibilité des mots CNC post-implantation ($r = 0,40$; $p = 0,0045$). Ce résultat, présenté figures 23.c et 23.d, est quasi-identique à ce qui a été observé dans l'étude d'Iowa (Rubinstein et al., 1999).

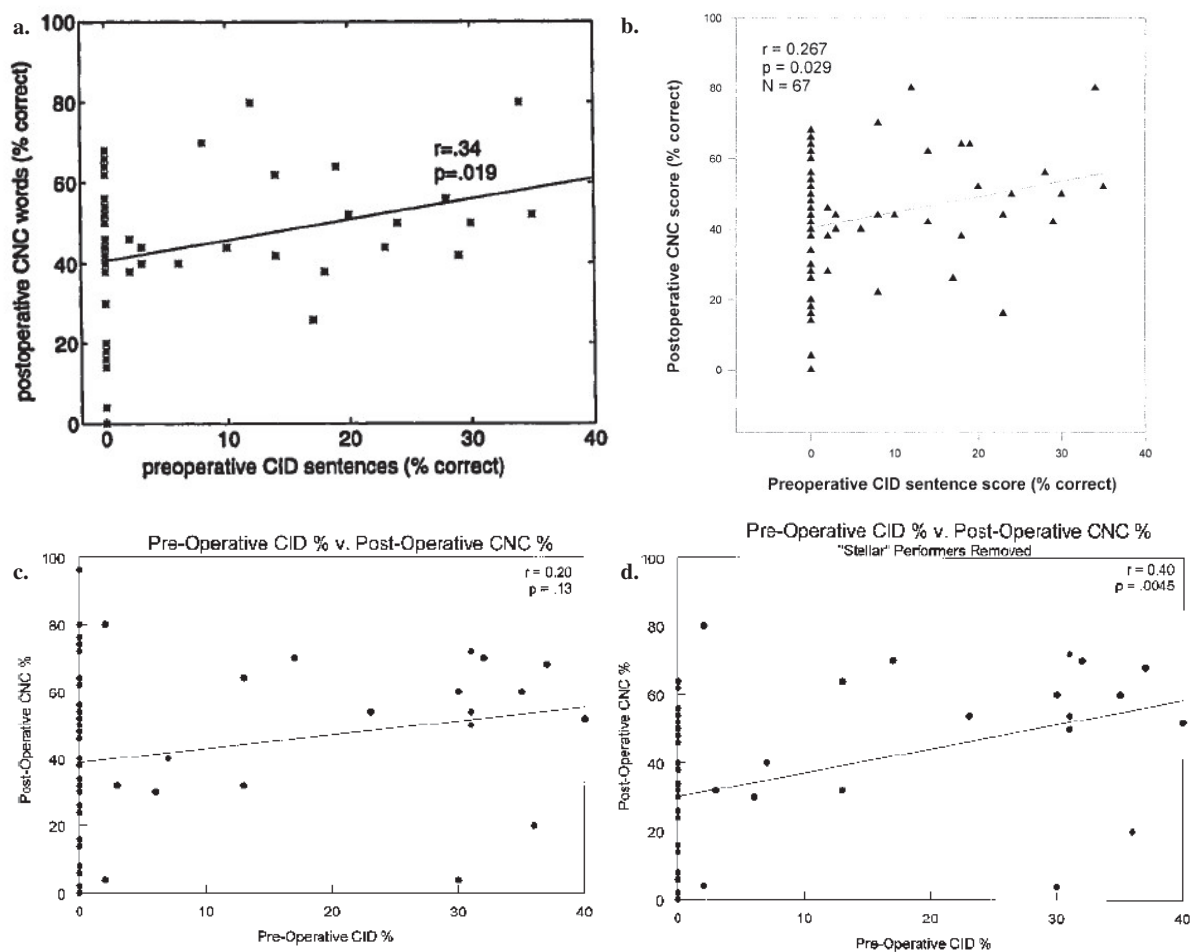


Figure 23. reconnaissance des mots post-implantation en fonction de la reconnaissance des phrases pré-implantation : a. d'après Rubinstein et al. (1999) chez 46 sujets implantés ; b. d'après Gomaa et al. (2003) chez 60 sujets ; c. et d. d'après Friedland et al. (2003), respectivement chez 58 sujets implantés (cadre de gauche) et chez 50 sujets qui n'étaient pas des « star patients » (cadre de droite).

Dans l'étude d'Adunka et al. (2008) décrite précédemment, les auteurs n'ont pas trouvé de différences de performances entre le groupe de patients qui a des restes auditifs et le groupe qui n'en a pas, quelle que soit la session de test et quel que soit le test utilisé (figure 18).

D'après Adunka et al., la présence d'audition résiduelle préopératoire ne fournirait pas un avantage pronostique notable pour les résultats de perception de la parole avec l'implant. La large disparité de performances entre les patients ne serait donc pas liée à l'audition résiduelle pré-implantation. Ces résultats sont en contradiction avec ce qui a été observés dans les travaux précédents qui ont investigué l'effet de divers facteurs sur les performances avec l'implant (Rubinstein et al., 1999 ; Gomaa et al., 2003 ; Friedland et al., 2003). Cependant

dans ces études, l'effet de l'audition résiduelle est significatif, mais les corrélations sont faibles. Dans l'étude de Friedland et al., lorsque tous les patients étaient pris en considération pour l'analyse, l'audition résiduelle ne montrait pas non plus d'effet significatif sur les performances. Les études n'ont pas mis en évidence, de façon unanime, que l'audition résiduelle pré-implantation était un bon facteur pronostique des performances post-implantation. Cependant cette audition résiduelle, si elle peut être exploitée acoustiquement après l'implantation (en ipsi- ou en contro-latéral), permet aux implantés de comprendre considérablement mieux, en particulier en milieu bruyant (Gantz et al., 2009). Cette stimulation bimodale proposée depuis une dizaine d'années s'appelle l'EAS (Electric-Acoustic Stimulation : Von Ilberg et al., 1999) et elle est actuellement en pleine expansion.

Gantz et al. (2009) ont montré que chez ces patients qui bénéficient de la stimulation bimodale, le score de perception de la parole pré-implantation est fortement relié au score post-implantation ($r = 0,39$; $p < 0,001$). Lorsque l'audition a pu être conservée et exploitée après la chirurgie, il apparaît que la compréhension dans le bruit est significativement meilleure et le SRT est corrélé au seuil de perception moyen à 125, 250 et 500 Hz (figure 24 : $r = 0,62$; $p < 0,05$). Pour la stimulation électro-acoustique, l'étude de Gantz et al. est d'ailleurs la seule à avoir évalué l'effet de la durée de surdité à ce jour. Selon eux, pour les sujets qui ne tirent pas de bénéfice de l'EAS ($n = 15$), 91% de la variance observée est expliquée par le score d'intelligibilité pré-implantation et par la durée de surdité dans les hautes fréquences. Pour les sujets qui obtiennent de bonnes performances avec l'EAS ($n = 53$), ces 2 facteurs expliquent 33% de la variance des résultats.

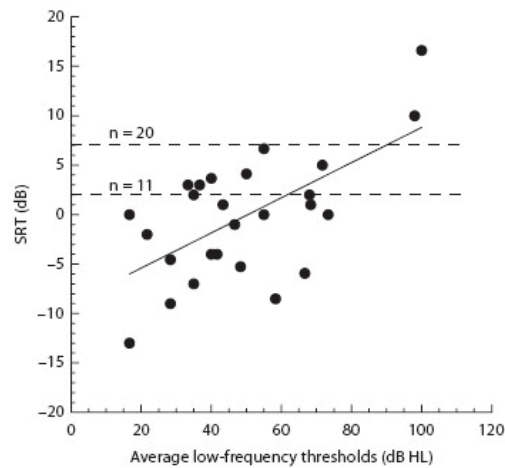


Figure 24. Seuil de reconnaissance de la parole pour des mots dissyllabiques et du bruit cocktail party, en fonction du seuil d'audition moyen dans les basses fréquences (à 125, 250 et 500 Hz). D'après Gantz et al. (2009).

Dans un autre laboratoire américain, Dorman et al. (2008) ont évalué le bénéfice fourni par une prothèse controlatérale à l'implant, en mesurant l'intelligibilité de la parole dans le silence et dans le bruit chez 65 sujets implantés unilatéraux (groupe « conventional ») versus 15 patients portant un implant d'un côté et une aide auditive du côté non implanté (groupe « EAS ») et ils ont comparé leurs données à celles des 55 sujets testés dans l'étude de Helms et al. en 1997 (groupe « Helms »).

Dans un premier temps, pour évaluer le bénéfice de la prothèse controlatérale, ils ont montré que leurs patients EAS en condition implant seul avaient des performances similaires à un échantillon aléatoire de sujets portant un implant cochléaire, pour la compréhension de mots CNC type consonne/noyau/consonne. Leurs 15 patients en condition implant seul obtiennent un score moyen de 54% et dans l'étude de Helms, les patients avaient un score moyen de 55%. A priori, il n'y a donc pas de différences de performances entre l'implanté moyen dans l'étude de 2008 et l'implanté moyen de 1997. De plus, la distribution des scores est sensiblement identique entre les 2 études. Par contre en condition EAS, les sujets EAS qui comprenaient 54% en condition implant seul voient leur intelligibilité s'améliorer de 19% lorsque l'implant est combiné à la prothèse controlatérale. Les résultats restent tout de même dispersés, mais cette dispersion est moins large qu'en condition implant seul (figure 25).

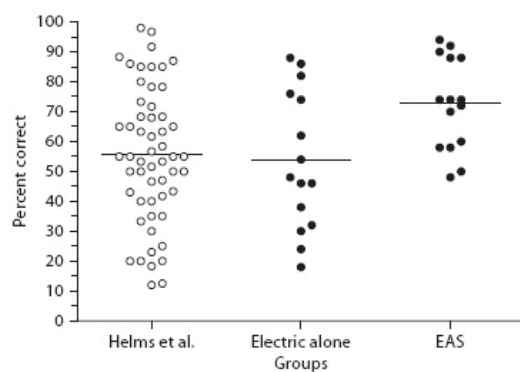


Figure 25. Intelligibilité des mots CNC de 15 patients implantés ayant de l'audition résiduelle du côté non implanté. Les scores sont présentés en condition implant seul, implant + prothèse controlatérale et la figure montre les données de l'étude de Helms et al. (1997) pour comparaison. D'après Dorman et al. (2008).

Dans un deuxième temps, Dorman et al. ont comparé les performances de leurs 65 sujets implantés unilatéraux aux performances de leurs 15 sujets EAS, pour différents tests de compréhension dans le silence et dans le bruit, parmi lesquels reconnaissance des mots CNC, reconnaissance des consonnes, des voyelles, des phrases dans le silence et dans le bruit à des RSB de +10 et +5 dB. Les résultats, présentés figure 26, ne semblent pas montrer une amélioration flagrante des performances chez les patients EAS par rapport aux meilleurs patients utilisateurs de l'implant conventionnel. Ce résultats est plutôt surprenant étant donné que l'information acoustique basse fréquence fournit des indices qui ne sont pas codés ou qui sont mal codés par l'implant cochléaire conventionnel. De plus, nous constatons qu'il existe toujours une large dispersion des résultats, ce qui suggère encore une fois que les performances obtenues avec l'implant sont patient dépendantes.

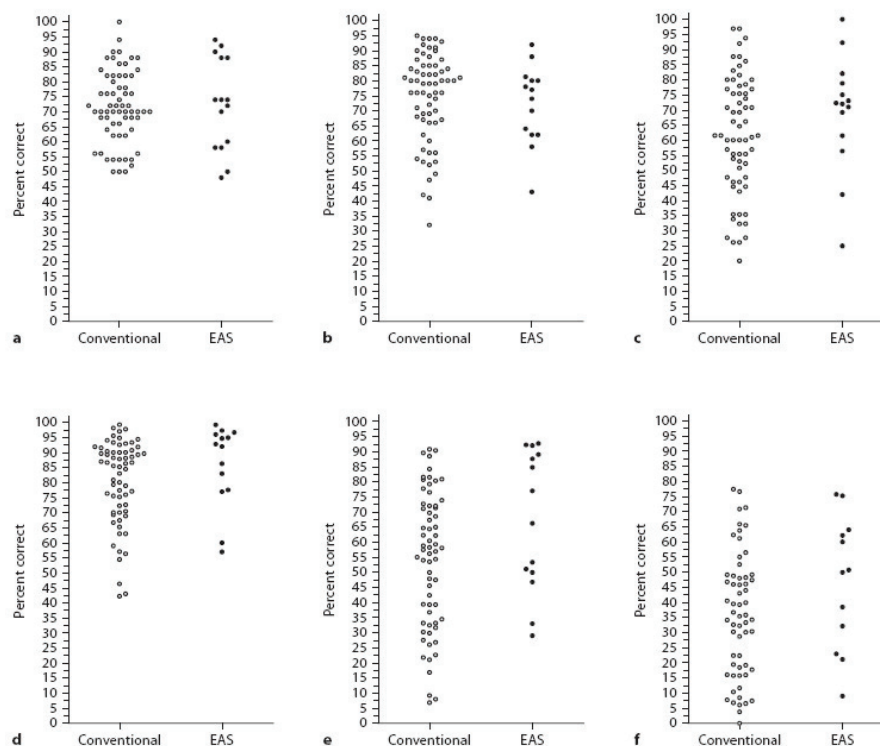


Figure 26. Pourcentage d'intelligibilité des patients implantés cochléaires conventionnels qui ont un score supérieur à 50% pour le test CNC (65 sujets) et des sujets EAS (15 sujets). Chaque point indique la performance d'un patient : a. mots CNC, b. consonnes, c. voyelles, d. phrases dans le silence, e. phrases dans le bruit à +10 dB de RSB, f phrases dans le bruit à +5 dB de RSB. D'après Dorman et al. (2008).

Conclusion

Cette brève revue de littérature montre que les performances des implantés cochléaires se sont considérablement améliorées depuis les années 80's (début de la commercialisation de l'implant cochléaire ; Zeng et al., 2008), en particulier grâce aux améliorations des stratégies de codage du son et à l'évolution des indications. Toutefois la technologie de l'implant, aussi évoluée soit-elle, ne pourra jamais fournir une intelligibilité parfaite, en particulier dans les conditions d'écoute difficile. D'une part il existe des différences de contraintes psychophysiques par rapport au normo-entendant, notamment pour la discrimination de fréquence et pour la discrimination d'intensité. D'autre part de nombreux facteurs tels que des longues durées de surdité altèrent sévèrement l'intégrité du système auditif, ce qui a pour conséquence des performances de compréhension limitées.

Références

- Abbas PJ, Brown CJ. Electrically evoked auditory brainstem response: growth of response with current level. *Hear Res.* 1991a;51:123-37.
- Abbas PJ, Brown CJ. Electrically evoked auditory brainstem response: refractory properties and strength-duration functions. *Hear Res.* 1991b;51:139-47.
- Adunka OF, Buss E, Clark MS, Pillsbury HC, Buchman CA. Effect of preoperative residual hearing on speech perception after cochlear implantation. *Laryngoscope.* 2008;118:2044-9.
- Blamey P, Arndt P, Bergeron F, Bredberg G, Brimacombe J, Facer G, Larky J, Lindström B, Nedzelski J, Peterson A, Shipp D, Staller S, Whitford L. Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants. *Audiol Neurotol.* 1996;1:293-306.
- Busby PA, Tong YC, Clark GM. Psychophysical studies using a multiple-electrode cochlear implant in patients who were deafened early in life. *Audiology.* 1992;31:95-111.
- Cooper HR. Selection criteria and prediction of outcomes. 2006. In: Cooper HR, Craddock LC (Eds.), *Cochlear Implants: A Practical Guide*, second ed. John Wiley & Sons, Hoboken, NJ, pp. 132-150.
- Dietrich V, Nieschalk M, Stoll W, Rajan R, Pantev C. Cortical reorganization in patients with high frequency cochlear hearing loss. *Hear Res.* 2001;158:95-101.
- Di Nardo W, Di Giuda D, Scarano E, Picciotti PM, Galla S, De Rossi G. Auditory cortical activation in severe-to-profound hearing-impaired patients monitored by SPET. *Acta Otorhinolaryngol Ital.* 2006;26:191-7.
- Dorman MF, Gifford RH, Spahr AJ, McKarns SA. The benefits of combining acoustic and electric stimulation for the recognition of speech, voice and melodies. *Audiol Neurotol.* 2008;13:105-12.
- Eshraghi AA, Frachet B, Van De Water TR, Eter E. Surdit  de l'adulte, r parer ou compenser. *Rev Prat.* 2009;59 :645-52.
- Friedland DR, Venick HS, Niparko JK. Choice of ear for cochlear implantation: the effect of history and residual hearing on predicted postoperative performance. *Otol Neurotol.* 2003;24:582-9.
- Fujiki N, Naito Y, Hirano S, Kojima H, Shiomi Y, Nishizawa S, Konishi J, Honjo I. Correlation between rCBF and speech perception in cochlear implant users. *Auris Nasus Larynx.* 1999;26:229-36.
- Gall go S, Frachet B, Micheyl C, Truy E, Collet L. Cochlear implant performance and electrically-evoked auditory brain-stem response characteristics. *Electroencephalogr Clin Neurophysiol.* 1998;108:521-5.
- Gantz BJ, Hansen MR, Turner CW, Oleson JJ, Reiss LA, Parkinson AJ. Hybrid 10 clinical trial: preliminary results. *Audiol Neurotol.* 2009;14 Suppl 1:32-8.

- Geier L, Barker M, Fisher L, Opie J. The effect of long-term deafness on speech recognition in postlingually deafened adult CLARION cochlear implant users. *Ann Otol Rhinol Laryngol Suppl.* 1999;177:80-3.
- Gomaa NA, Rubinstein JT, Lowder MW, Tyler RS, Gantz BJ. Residual speech perception and cochlear implant performance in postlingually deafened adults. *Ear Hear.* 2003;24:539-44.
- Green KM, Bhatt YM, Mawman DJ, O'Driscoll MP, Saeed SR, Ramsden RT, Green MW. Predictors of audiological outcome following cochlear implantation in adults. *Cochlear Implants Int.* 2007;8:1-11.
- Green KM, Julyan PJ, Hastings DL, Ramsden RT. Auditory cortical activation and speech perception in cochlear implant users: effects of implant experience and duration of deafness. *Hear Res.* 2005;205:184-92.
- Groenen PA, Makhdom M, van den Brink JL, Stollman MH, Snik AF, van den Broek P. The relation between electric auditory brain stem and cognitive responses and speech perception in cochlear implant users. *Acta Otolaryngol.* 1996;116:785-90.
- Haumann S, Lenarz T, Büchner A. Speech perception with cochlear implants as measured using a roving-level adaptive test method. *ORL J Otorhinolaryngol Relat Spec.* 2010;72:312-8.
- Helms J, Müller J, Schön F, Moser L, Arnold W, Janssen T, Ramsden R, von Ilberg C, Kiefer J, Pfennigdorf T, Gstöttner W, Baumgartner W, Ehrenberger K, Skarzynski H, Ribari O, Thumfart W, Stephan K, Mann W, Heinemann M, Zorowka P, Lippert KL, Zenner HP, Bohndord M, Hüttenbrink K, Müller-Aschoff E, Hofmann G, Freigang B, Begall K, Ziese M, Forgbert O, Häusler R, Vischer M, Schlatter T, Schlöndorff G, Korves B, Döring H, Gerhardt HJ, Wagner H, Schorn K, Schilling V, Baumann U, Kastenbauer E, Albegger K, Mair A, Gammert Ch, Mathis A, Streitberger Ch, Hochmair-Desoyer I. Evaluation of performance with the COMBI40 cochlear implant in adults: a multicentric clinical study. *ORL J Otorhinolaryngol Relat Spec.* 1997;59:23-35.
- Guiraud J, Gallego S, Arnold L, Boyle P, Truy E, Collet L. Effects of auditory pathway anatomy and deafness characteristics? (1): On electrically evoked auditory brainstem responses. *Hear Res.* 2007a;223:48-60.
- Guiraud J, Gallego S, Arnold L, Boyle P, Truy E, Collet L. Effects of auditory pathway anatomy and deafness characteristics? Part 2: On electrically evoked late auditory responses. *Hear Res.* 2007b;228:44-57.
- Kakigi A, Hirakawa H, Harel N, Mount RJ, Harrison RV. Tonotopic mapping in auditory cortex of the adult chinchilla with amikacin-induced cochlear lesions. *Audiology.* 2000;39:153-60.
- Kessler DK, Loeb GE, Barker MJ. Distribution of speech recognition results with the Clarion cochlear prosthesis. *Ann Otol Rhinol Laryngol Suppl.* 1995;166:283-5.
- Kral A, Tillein J. Brain plasticity under cochlear implant stimulation. *Adv Otorhinolaryngol.* 2006;64:89-108.

- Krueger B, Joseph G, Rost U, Strauss-Schier A, Lenarz T, Buechner A. Performance groups in adult cochlear implant users: speech perception results from 1984 until today. *Otol Neurotol*. 2008;29:509-12.
- Lee DS, Lee JS, Oh SH, Kim SK, Kim JW, Chung JK, Lee MC, Kim CS. Cross-modal plasticity and cochlear implants. *Nature*. 2001;409(6817):149-50.
- Leung J, Wang NY, Yeagle JD, Chinnici J, Bowditch S, Francis HW, Niparko JK. Predictive models for cochlear implantation in elderly candidates. *Arch Otolaryngol Head Neck Surg*. 2005;131:1049-54.
- Makhdoun MJ, Groenen PA, Snik AF, van den Broek P. Intra- and interindividual correlations between auditory evoked potentials and speech perception in cochlear implant users. *Scand Audiol*. 1998;27:13-20.
- Maurer J, Collet L, Pelster H, Truy E, Gallégo S. Auditory late cortical response and speech recognition in Digisonic cochlear implant users. *Laryngoscope*. 2002;112:2220-4.
- McDermott HJ, Lech M, Kornblum MS, Irvine DR. Loudness perception and frequency discrimination in subjects with steeply sloping hearing loss: possible correlates of neural plasticity. *J Acoust Soc Am*. 1998;104:2314-25.
- Robertson D, Irvine DR. Plasticity of frequency organization in auditory cortex of guinea pigs with partial unilateral deafness. *J Comp Neurol*. 1989;282:456-71.
- Rubinstein JT, Parkinson WS, Tyler RS, Gantz BJ. Residual speech recognition and cochlear implant performance: effects of implantation criteria. *Am J Otol*. 1999;20:445-52.
- Schwaber MK, Garraghty PE, Kaas JH. Neuroplasticity of the adult primate auditory cortex following cochlear hearing loss. *Am J Otol*. 1993;14:252-8.
- Shipp DB, Nedzelski JM. Prognostic indicators of speech recognition performance in adult cochlear implant users: a prospective analysis. *Ann Otol Rhinol Laryngol Suppl*. 1995;166:194-6.
- Thai-Van H, Micheyl C, Moore BC, Collet L. Enhanced frequency discrimination near the hearing loss cut-off: a consequence of central auditory plasticity induced by cochlear damage? *Brain*. 2003;126:2235-45.
- Thai-Van H, Micheyl C, Norena A, Collet L. Local improvement in auditory frequency discrimination is associated with hearing-loss slope in subjects with cochlear damage. *Brain*. 2002;125:524-37.
- Truy E, Bouccara D, Sterkers O, Triglia JM. Techniques chirurgicales d'implantation d'aides auditives en otoneurologie. EMC (Elsevier Masson SAS, Paris), Techniques chirurgicales – Tête et cou, 2009 ;46-038:1-20.
- Tyler RS, Parkinson AJ, Wilson BS, Witt S, Preece JP, Noble W. Patients utilizing a hearing aid and a cochlear implant: speech perception and localization. *Ear Hear*. 2002;23:98-105.
- Tyler RS, Parkinson AJ, Woodworth GG, Lowder MW, Gantz BJ. Performance over time of adult patients using the Ineraid or nucleus cochlear implant. *J Acoust Soc Am*. 1997;102:508-22.

- UK Cochlear Implant Study Group. Criteria of candidacy for unilateral cochlear implantation in postlingually deafened adults I: theory and measures of effectiveness. *Ear Hear.* 2004;25:310-35.
- UK Cochlear Implant Study Group. Criteria of candidacy for unilateral cochlear implantation in postlingually deafened adults II: cost-effectiveness analysis. *Ear Hear.* 2004;25:336-60.
- UK Cochlear Implant Study Group. Criteria of candidacy for unilateral cochlear implantation in postlingually deafened adults III: prospective evaluation of an actuarial approach to defining a criterion. *Ear Hear.* 2004;25:361-74.
- van Dijk JE, van Olphen AF, Langereis MC, Mens LH, Brokx JP, Smoorenburg GF. Predictors of cochlear implant performance. *Audiology.* 1999;38:109-16.
- von Ilberg C, Kiefer J, Tillein J, Pfenningdorff T, Hartmann R, Stürzebecher E, Klinke R. Electric-acoustic stimulation of the auditory system. New technology for severe hearing loss. *ORL J Otorhinolaryngol Relat Spec.* 1999;61:334-40.
- Wilson BS. Speech processing strategies. 2006. In: Cooper HR, Craddock LC (Eds.), *Cochlear Implants: A Practical Guide*, second ed. John Wiley & Sons, Hoboken, NJ, pp. 21-69.
- Wilson BS, Dorman MF. Cochlear implants: a remarkable past and a brilliant future. *Hear Res.* 2008;242:3-21.
- Zeng FG. Trends in cochlear implants. *Trends Amplif.* 2004;8:1-34.
- Zeng FG, Rebscher S, Harrison WV, Sun X, Feng H. Cochlear Implants: System Design, Integration and Evaluation. *IEEE Rev Biomed Eng.* 2008 Jan 1;1:115-142.

I.7. Cas particulier des sujets qui ont des restes sur les graves

Pour l'appareillage des surdités partielles, de nombreuses études montrent que l'amplification des hautes fréquences est inefficace et qu'elle peut parfois nuire à la compréhension. C'est pourquoi dans certains cas il peut être recommandé de ne pas les amplifier. Certains audiologistes recommandent l'utilisation du TEN test (Threshold Equalizing Noise, Moore et al., 2000) pour détecter d'éventuelles zones inertes cochléaires (ZIC) et évaluer si une zone de fréquences doit être amplifiée ou non (Vickers et al., 2001 ; Baer et al., 2002). Mais il semble que ce test additionnel n'est clairement pas nécessaire pour décider si les hautes fréquences doivent être amplifiées. Et d'après les cliniciens expérimentés, pour l'appareillage des patients qui présentent ce type de perte, il suffirait tout simplement de ne pas fournir d'amplification sur les zones où la perte est supérieure à 90 dB HL (Summers, 2004).

Depuis la fin des années 90's, de nombreux travaux ont été publiés sur l'intérêt d'amplifier ou non les hautes fréquences chez les sourds partiels. Nous trouvons en particulier les travaux des équipes de Christopher Turner à Iowa (Hogan et Turner, 1998 ; Turner et Cummings, 1999) et de Brian Moore à Cambridge (Vickers et al., 2001 ; Baer et al., 2002). Les résultats et les interprétations de ces différentes études suscitent toujours un grand intérêt aujourd'hui et sont encore très discutés.

En 1998, Hogan et Turner (1998) avaient réalisé des tests de compréhension de parole filtrée passe-bas à différentes fréquences de coupure chez 5 normo-entendants et 9 malentendants avec une surdité partielle. Pour chaque condition, les auteurs quantifiaient l'audibilité en utilisant l'index d'articulation. Ensuite, pour les 2 populations testées, ils évaluaient le bénéfice en terme de compréhension fourni par l'ajout d'une bande additionnelle 1/3 d'octave de plus haute fréquence.

Pour l'interprétation de leurs résultats, ils ont introduit la notion d'efficacité pour définir une mesure qui permet d'apprécier comment les malentendants utilisent l'information de parole, à des niveaux audibles et à différentes fréquences, par rapport aux sujets normo-entendants.

L'équation de l'efficacité est la suivante :

$$\text{efficiency} = \frac{\Delta \text{ score} / \Delta \text{ AI(hearing-impaired)}}{\Delta \text{ score} / \Delta \text{ AI(normal group)}}$$

Si l'efficacité a une valeur de 1, cela signifie que le malentendant utilise l'information additionnelle de la même façon que le normo-entendant.

Si l'efficacité est supérieure à 1, c'est que le malentendant utilise mieux l'information additionnelle que le normo-entendant.

Si l'efficacité est entre 0 et 1, c'est que le bénéfice fourni par l'ajout d'information audible est moins important pour le malentendant que pour le normo-entendant.

Enfin, si l'efficacité est inférieure à 0, cela signifie que l'ajout d'information audible est contreproductif, c'est-à-dire que fournir de l'audibilité sera délétère.

Leurs résultats montrent que pour les plus basses fréquences (1000 et 1600 Hz) qui ne sont que légèrement altérées chez les malentendants, les efficacités sont agglomérées autour d'une valeur unité, où les sujets ont généralement une audition normale. Dans les plus hautes fréquences où les sujets ont un degré de surdité plus important, l'efficacité se dégrade et atteint des valeurs négatives, en particulier dès que la perte devient supérieure à 55 dB HL.

Cet effet se confirme lorsque l'efficacité est représentée en fonction du seuil d'audition pour les différentes fréquences. Les résultats suggèrent qu'au-delà de 55 dB de perte, la restauration des informations haute-fréquence risque de nuire à la compréhension. Dans ce cas, les auteurs considèrent que l'amplification est contreproductive.

La même année, Ching et al. (1998) ont examiné la relation entre l'audibilité et l'intelligibilité de la parole pour une cinquantaine de sujets avec différentes configurations de pertes auditives. Les auteurs ont mis en évidence que, pour les patients qui présentent une perte sévère à profonde ou une perte à pente abrupte, les malentendants comprennent moins bien que ce qui peut être espéré sur la base de l'audibilité. Un de leur modèles montre également que, pour les sujets dont le seuil auditif à 4 kHz est supérieur à 80 dB HL, le fait de rendre audible l'information dans la bande 2,8–5,6 kHz ne contribue pas à l'intelligibilité. Pour ces sujets, les auteurs recommandent de ne pas amplifier les composantes haute-fréquence de la parole. Selon eux les sujets ayant un tel degré de surdité ne sont pas capables d'extraire l'information d'un signal qui serait rendu audible dans cette région.

En 1999, Turner et Cummings ont évalué l'amélioration des performances en fonction de l'augmentation d'audibilité chez 16 sujets : 5 étaient normo-entendants, 10 étaient malentendants avec une surdité prédominant dans les aigus et 1 sujet avait une perte plate. Dans leur expérimentation, les auteurs augmentent l'audibilité en augmentant l'intensité de stimulation. L'idée de ces mesures est de fournir de plus en plus d'énergie de parole au dessus du seuil de perception, pour que les indices de parole localisés dans les zones fréquentielles lésées soient de plus en plus perceptibles, fournissant ainsi de plus en plus d'information.

Pour les patients qui ont une surdité partielle, ces tests montrent que, jusqu'à une certaine limite, l'élévation du niveau de présentation résulte en une augmentation de la compréhension. Mais au-delà de cette limite, propre à chaque sujet, une augmentation supplémentaire du niveau de présentation (= de l'audibilité) ne va pas améliorer les performances et elle peut dans certains cas les dégrader (figure 11).

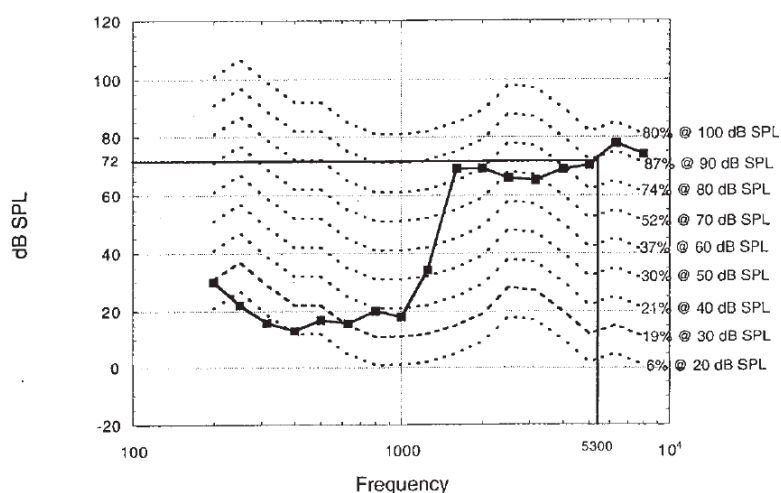


Figure 11. La courbe en traits pleins représente les seuils audiométriques pour un sujet présentant une perte auditive dans les hautes fréquences. Les courbes en pointillés présentent le niveau pic des consonnes pour de la parole présentée à neuf intensités différentes (indiquées à droite du graphique). A la droite de chaque spectre de parole sont indiqués les pourcentages de reconnaissance pour les différents niveaux de présentation. D'après Turner et Cummings (1999).

Les auteurs de ces trois études recommandent donc de ne pas amplifier les hautes fréquences lorsque la perte dépasse un certain degré, à savoir 55-60 dB HL. Comment expliquer les résultats observés ?

Zones inertes cochléaires et parole filtrée passe-bas

Suite aux études décrites précédemment, Vickers et al. (2001) et Baer et al. (2002) se sont interrogés sur les limites utiles de l'amplification des hautes fréquences, ainsi que sur l'effet de l'audibilité des hautes fréquences sur l'intelligibilité, en présence de ZIC.

Leurs études consistent à soumettre les sujets malentendants atteints de surdité neurosensorielle dans les hautes fréquences à des tests de parole filtrée passe-bas (dans le silence pour Vickers et dans le bruit pour Baer) et de corréler les différents patterns de réponses avec la présence ou non de ZIC. Leurs résultats montrent que si le patient n'a pas de ZIC, la compréhension augmente lorsque la fréquence de coupure du filtre passe-bas augmente, jusqu'à 7,5 kHz. Si par contre le patient présente une ZIC, le pattern de réponse « reconnaissance en fonction de la fréquence de coupure passe-bas » va soit augmenter au début et stagner à un plafond, soit augmenter au début puis redescendre. Ce qui signifie que les patients qui ont des ZIC sont moins capables d'utiliser l'information haute-fréquence amplifiée que des patients qui n'ont pas de ZIC. Les auteurs estiment qu'en général, chez les patients qui ont une ZIC le maximum de performance est atteint lorsque la fréquence de coupure augmente jusqu'à une valeur de 50% à 100% au-delà de la fréquence estimée de la zone morte.

Les observations de Vickers et de Baer sont en accord avec ce qui a été montré auparavant et ils sont particulièrement intéressants dans la mesure où ils fournissent une recommandation nouvelle au clinicien, à savoir que dans certains cas il ne faut pas amplifier les fréquences qui sont 1,5 à 2 fois supérieures à la fréquence de la zone morte. Leurs travaux sont toutefois critiquables dans la mesure où leurs patients qui présentaient des ZICs avaient des pertes beaucoup plus importantes que leurs patients qui n'avaient pas de ZIC. Il n'est donc pas possible d'apprécier si l'incapacité des patients à utiliser les hautes fréquences est due à la présence de ZIC, à un problème d'audibilité, ou bien tout simplement à la sévérité de la perte.

Ces résultats ont plus tard été remis en question par Simpson et al. (2005). De manière analogue à Vickers et al (2001), les auteurs ont réalisé des tests d'audiométries tonales normales et approfondies avec le TEN test, ainsi que des audiométries vocales avec de la parole filtrée passe-bas. Selon eux la présence de ZIC n'affecterait pas l'utilisation des hautes fréquences pour comprendre et leurs sujets testés seraient capables d'utiliser cette information haute fréquence lorsqu'ils y ont accès.

D'après Moore (2001), lorsque la perte auditive dépasse 90 dB HL dans les hautes fréquences, alors elle peut être communément associée à une ZIC. On peut alors s'interroger sur la nécessité de réaliser le test des ZIC pour prédire si l'amplification des hautes fréquences va fournir un bénéfice. Certains chercheurs ont d'ailleurs noté que la présence de ZIC et la sévérité de la perte dans une certaine région fréquentielle covariaient (Rankovic, 2002 ; Summers, 2004). On peut tout simplement supposer qu'au-delà d'un certain degré de perte (dès 55 dB HL d'après Hogan et Turner–1998, jusqu'à 90 dB HL d'après Summers–2004), il n'est pas utile d'amplifier les hautes fréquences.

Amos et Humes, en 2007, ont observé des résultats similaires à ceux de Vickers et de Baer. Dans leur protocole d'étude ils faisaient écouter de la parole filtrée passe bas en bande étroite (200–1600 Hz), en bande medium (200–3200 Hz) et en bande large (200–6400 Hz), dans le silence (RSB = +20 dB) et dans le bruit (RSB = +5 dB) à 36 sujets (3 groupes) malentendants plus ou moins âgés ayant divers degrés de surdité dans les hautes fréquences, mais qui ne présentaient pas forcément de ZIC (le TEN test n'a pas été utilisé dans cette étude). Les résultats ne révèlent aucune amélioration des performances lorsque la bande passante du filtre passe de medium à large, ce qui laisse supposer que l'information au-delà de 3200 Hz ne contribue pas à l'intelligibilité de la parole pour ces patients.

A la même époque, plusieurs équipes dont Hornsby et Ricketts (2006) remettent en question l'idée qu'il ne faut pas amplifier les hautes-fréquences chez ces malentendants. Les auteurs ont conduit leur étude chez des sujets avec une surdité partielle dans les aigus et chez un groupe de malentendants avec une perte plate qu'ils avaient testés dans une précédente étude (Hornsby et Ricketts, 2003). Les sujets étaient soumis à des tests de reconnaissance de parole filtrée passe-haut et passe-bas. Contrairement à ce qui a été observé par Vickers et al. (2001), les sujets testés ont des performances qui augmentent de façon quasi monotone lorsque la fréquence de coupure passe-bas augmente, et ce quelque soit la configuration de la perte. Le score ne diminue jamais en condition large bande. C'est donc que les sujets utilisent a priori l'information haute-fréquence lorsqu'ils y ont accès, mais ils sous-exploitent ces hautes fréquences au profit des basses fréquences, d'où un décalage de la fréquence d'intersection vers les basses fréquences. Ces résultats suggèrent donc que l'audibilité des hautes fréquences dans les zones où la perte est importante (jusqu'à 90 dB HL), ne nuit pas forcément à la compréhension.

Pour mieux comprendre si la dégradation des performances en condition large bande était spécifique au sourd partiel, Horwitz et al. (2008) ont proposé d'évaluer dans quelle mesure l'amplification des hautes fréquences pouvait aider ou nuire à la compréhension, en fonction de la perte auditive ou des facteurs de corrections proposés aux patients. Comme pour les études décrites précédemment, les auteurs utilisaient des tests d'audiométrie vocale filtrée passe-bas à différentes fréquences de coupure. L'étude était menée chez 16 malentendants et les auteurs ont inclus 18 normo-entendants afin d'obtenir des données normatives. De manière globale, l'intelligibilité augmente lorsque la fréquence de coupure passe-bas augmente. Les scores prédits sont inférieurs aux scores observés pour les normo-entendants tandis qu'ils sont supérieurs aux scores observés pour les malentendants. Pour 10 des 16 malentendants, le plus haut score est obtenu pour une bande-passante qui n'est pas maximum, généralement 4500 Hz ; il y a donc dans certains cas une diminution des performances en condition large bande. Cependant la même tendance se retrouve chez 7/18 normo-entendants et la différence de performance entre le score le plus haut et le score large bande est généralement très faible.

Dans l'étude de Horwitz et al., La saturation des performances au-delà d'une certaine fréquence de coupure se retrouve chez 40% des normo-entendants. Cela nous conduit à penser que ce pattern de réponse n'est pas nécessairement lié à la surdité. D'après les auteurs, pour les pertes auditives jusqu'à 80 dB dans les hautes fréquences, les variations de performances seraient indépendantes du degré de perte auditive et contrairement aux suggestions de Hogan et Turner (1998), leurs résultats ne mettent pas en évidence que l'amplification peut être contre-productive au-delà d'une certaine perte. Toutefois les malentendants seraient plus efficaces pour extraire l'information à laquelle ils ont été habitués, dans la zone fréquentielle où leur audition est la meilleure ; un constat semblable avait été fait par Vestergaard en 2003.

Limitations techniques des appareils auditifs et alternatives

Les premières études utilisant du filtrage passe-bas chez le sourd partiel (Hogan et Turner, 1998 ; Ching et al., 1998 ; Vickers et al., 2001 ; Baer et al., 2002) s'accordent à dire que les sourds partiels n'utilisent pas les hautes fréquences pour comprendre et préconisent de ne pas amplifier ces fréquences pour améliorer la compréhension des patients. Cependant, des études plus récentes utilisant des populations et des procédures de tests similaires (Mackersie et al., 2004 ; Simpson et al., 2005 ; Hornsby et Ricketts, 2006 ; Horwitz et al., 2008) semblent montrer que les patients sont capables d'exploiter l'information haute fréquence lorsqu'elle est fournie.

On peut supposer qu'à l'époque les méthodes de réglages étaient inadaptées et les technologies étaient trop limitées pour réhabiliter les surdités partielles. Les hautes fréquences étaient probablement moins amplifiées à cause de la sensibilité au larsen et les patients avaient peut-être appris à comprendre sans les hautes fréquences qui n'étaient pas suffisamment audibles à l'époque, ce qui serait concordant avec le décalage de la fréquence d'intersection observé chez Hornsby et Ricketts (2006).

Les premières études que nous avons décrites ont été publiées à la fin des années 90's et au début des années 2000's. Or à cette époque la plupart des appareils auditifs étaient analogiques et les appareils numériques étaient peu répandus, les possibilités techniques étaient donc beaucoup plus limitées par rapport aux appareils actuellement disponibles. Les deux principaux problèmes pour l'appareillage de ces pertes étaient et sont d'ailleurs toujours les phénomènes de larsen acoustique et d'occlusion (Boeheim et al., 2010). Le larsen acoustique se manifeste à cause des fuites du signal sonore amplifié, qui s'échappent autour de l'embout et qui sont réamplifiées par l'aide auditive. Plus l'amplification est importante, plus le larsen est susceptible de se produire. Et pour les sourds partiels il est nécessaire d'appliquer un gain important dans les hautes fréquences. Ce phénomène de larsen peut être diminué en utilisant un embout étanche qui va obstruer le conduit auditif externe pour limiter les fuites, mais bien souvent de tels embouts sur de telles pertes auditives résultent en un effet d'occlusion qui va gêner le patient tout autant que le larsen. L'effet d'occlusion se produit lorsque le patient trouve que sa propre voix sonne anormalement. Les sons de parole prononcés par le patient, transmis par conduction osseuse, vont être piégés entre l'embout et la membrane tympanique. Ceci va résulter en une amplification des sons basse-fréquence (< 500 Hz) jusqu'à 20 dB et le patient va avoir une sensation de résonance. Une solution pour limiter ce phénomène consiste à mettre un évent dans l'embout pour aérer l'oreille, mais en contrepartie cet évent va favoriser l'apparition du larsen. C'est pourquoi l'appareillage des sourds partiels est assez délicat. Rétrospectivement, et en gardant à l'esprit que les appareils auditifs étaient encore analogiques il y a 10 ou 15 ans, nous pouvons supposer que ces deux phénomènes sont les principales causes pour lesquelles les auteurs de la fin des années 90's recommandaient de ne pas amplifier les hautes fréquences.

Parallèlement, à la même époque, l'implant d'oreille moyenne (IOM) tel que le Soundbridge de Symphonix s'est trouvé être une alternative à l'appareillage acoustique pour restituer plus efficacement les hautes fréquences. D'une part le patient n'avait plus d'effet d'occlusion car il n'avait pas d'embout dans l'oreille. D'autre part l'appareil était moins sensible au larsen du fait que le transducteur (électromagnétique) est éloigné du microphone qui capte les sons. D'après les publications qui comparent les performances de l'aide auditive à celles du Soundbridge, il apparaît clairement qu'au niveau du gain fonctionnel (différence entre seuil tonal non appareillé et seuil appareillé, en champ libre) l'IOM fournit des résultats identiques ou supérieurs à l'aide auditive, mais jamais moins bons (Frayssé et al., 2001 ; Luetje et al., 2002 ; Todt et al., 2002 ; Uziel et al., 2003 ; Truy et al., 2008 ; Boeheim et al., 2010 ; Tysome et al., 2010). Ces études présentent également les résultats des patients pour des tests d'audiométries vocales et montrent là encore la supériorité de l'implant Soundbridge par rapport à l'aide auditive conventionnelle. Si l'IOM fournit un meilleur gain fonctionnel dans les hautes fréquences et une meilleure intelligibilité, on peut supposer que les patients sourds partiels peuvent traiter l'information haute fréquence lorsqu'elle est audible.

La contribution des hautes fréquences pour l'intelligibilité de la parole chez les sourds partiels est très controversée. Certains pensent que ces fréquences ne sont pas utiles, d'autres suggèrent de les amplifier quand l'appareil l'autorise, d'autre enfin supposent que les malentendants sous-exploitent l'information haute fréquence mais l'utilisent quand même pour comprendre. La revue de littérature et la chronologie des publications laisse supposer que les recommandations cliniques des différents articles étaient probablement liées à la technologie des appareils de l'époque. Nous constatons également une large variabilité inter-individuelles dans ces études, pour les tests d'audiométries tonales et vocales et pour l'utilisation des différentes bandes fréquentielles. Il n'est donc pas exclu que la variabilité des résultats soit expliquée par des facteurs propres aux patients tels que l'étiologie de la surdité ou la présence de ZIC. Pour des explications alternatives sur l'effet délétère de l'amplification des hautes fréquences, nous renvoyons le lecteur à l'article 1 de la présente introduction.

Face à de tels patients, et si les patients ne souhaitent pas se faire implanter, le clinicien peut toujours proposer un appareillage acoustique à compression fréquentielle. Ces systèmes supposent que l'enveloppe temporelle soit préservée et le malentendant perçoit les sons de

façon plus grave et un peu plus naturelle. Selon McDermott et al. (1998) et Thai-Van et al. (2002 et 2003), ces patients avec une perte importante dans les hautes fréquences auraient une meilleure discrimination fréquentielle et une sur-représentation des neurones, au niveau cortical, à la fréquence de coupure de la perte (début de la perte). D'un point de vue théorique, le malentendant devrait donc être capable d'intégrer l'information transposée par les systèmes à compression fréquentielle. C'est ce que nous proposons d'étudier dans la section suivante

I.8. Appareillage des surdités dans les hautes fréquences par les techniques de réduction fréquentielle

Appareillage des surdités dans les hautes fréquences par les techniques de réduction fréquentielle

(sous presse à l’Ouïe magazine)

Fabien Seldran (1,2), Doctorant en Ingénierie Biomédicale, audioprothésiste.

Lionel Collet (1), Professeur des universités-Praticien Hospitalier, Chef du service d’audiologie et d’explorations oro-faciales des Hospices Civils de Lyon, Directeur de l’école d’Audioprothèse de Lyon.

Stéphane Gallégo (1,3), Docteur en Ingénierie Biomédicale, audioprothésiste.

(1) Hôpital Edouard Herriot, Pavillon U, service d’audiologie et d’explorations oro-faciales, Lyon

(2) Vibant Med-El, Sophia-Antipolis

(3) Audition Conseil, Lyon

Pour les malentendants qui perçoivent encore les sons graves mais qui ont une perte auditive profonde dans les fréquences aiguës, il est difficile pour le clinicien de prédire le type d’appareillage le plus adapté pour restituer les hautes fréquences (HF) de façon optimale. Le réflexe serait de dire « amplifions les hautes fréquences ! », mais malheureusement tout n’est pas aussi simple et deux problèmes majeurs limitent le bénéfice de l’amplification des hautes fréquences par les appareils conventionnels.

D’un point de vue technique le premier problème est lié au phénomène de larsen acoustique qui se produit lorsque l’amplification dans les aigus est trop importante. Le larsen acoustique se manifeste en raison des fuites du signal sonore amplifié, qui s’échappent autour de l’embout et qui sont ré-amplifiées par l’aide auditive. Plus l’amplification est importante, plus le larsen est susceptible de se produire. Et pour les surdités partielles il est nécessaire d’appliquer un gain important dans les hautes fréquences. Ce phénomène de larsen peut être diminué en utilisant un embout étanche qui obstrue le conduit auditif externe pour limiter les fuites, mais bien souvent de tels embouts sur de telles pertes auditives provoquent un effet d’occlusion qui gêne le patient tout autant que le larsen.

Le second problème se situe au niveau perceptif, en particulier si le patient présente des zones mortes cochléaires. Le terme de « Zones mortes Cochléaires » a été introduit en 2000 par Moore, pour qualifier une région de la cochlée dans laquelle les CCI et/ou les neurones fonctionnent tellement peu/mal qu'un son pur produisant un pic de vibration dans cette région est détecté par des cellules répondant habituellement à une autre fréquence caractéristique. Le son est détecté à un endroit où la vibration de la membrane basilaire sera moins importante, mais où les CCI et les neurones fonctionnent plus efficacement. Selon Moore (2001), une zone où la perte est supérieure à 90 dB HL dans les hautes fréquences, peut-être associée à une zone morte cochléaire. L'expression de zone inerte cochléaire (ZIC) est préférée par certaines équipes. Pour de tels degrés de perte, l'amplification des hautes fréquences risque de créer des distorsions au point d'interférer avec les sons perçus par les zones cochléaires encore saines. Dans ces situations, de nombreux travaux ont montré que l'amplification est souvent inefficace et le fait de rendre audibles les hautes fréquences peut dans certains cas avoir des effets délétères, en particulier dès que la perte auditive dépasse 55 à 60 dB (Ching et al., 1998 ; Hogan & Turner, 1998 ; Ching et al., 2001 ; Vickers et al., 2001 ; Turner, 2006). Comment expliquer ce phénomène ? Et comment y remédier ?

Les aides auditives conventionnelles étant limitées pour fournir des hautes fréquences à ces sujets, des méthodes alternatives de traitement du signal proposent de présenter l'information haute-fréquence de la parole dans des régions fréquentielles plus basses. Il existe différentes techniques de diminution fréquentielle, mais quelle que soit la méthode proposée, l'objectif reste le même : restaurer l'information HF non perceptible par ces sujets.

Depuis plusieurs décennies, on s'interroge sur l'intérêt d'utiliser des algorithmes à transposition fréquentielle pour restituer les hautes fréquences, mais jusque là ces algorithmes étaient difficiles à intégrer dans des systèmes embraqués types contour d'oreille. Ces dernières années les techniques de traitements numériques des signaux se sont considérablement développées. Il est désormais possible d'intégrer la compression fréquentielle dans les aides auditives tout en préservant la qualité naturelle des sons, grâce à des algorithmes tels que l'Audibility Extender de Widex (figure 1) ou le SoundRecover de Phonak (figure 2). Parmi les techniques de réduction fréquentielle (frequency lowering) proposées par les différents chercheurs et ingénieurs, citons la transposition fréquentielle et la compression linéaire ou non linéaire de fréquence. Depuis les années 1960-1970, de nombreuses équipes ont tenté d'adapter ces techniques mais les résultats avec ces systèmes

étaient en général assez mitigés. Dans cet article, nous présentons les différentes techniques, leurs principes, leurs limitations et les résultats qu'on peut en espérer.

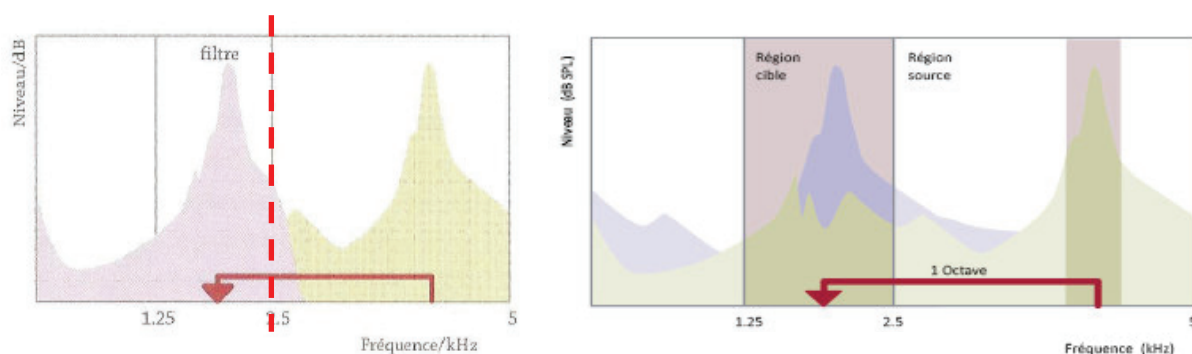


Figure 1. Principe de la transposition fréquentielle avec chevauchement par le système Audibility Extender de Widex. (Source : Widex)

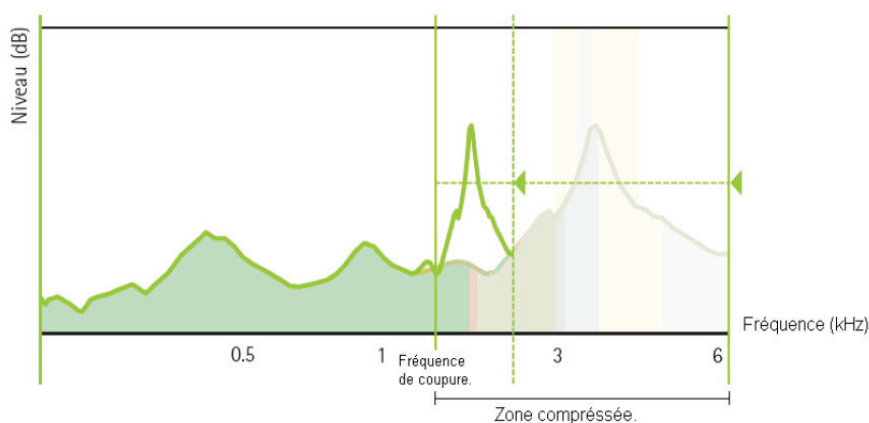


Figure 2. Principe de la compression non linéaire de fréquence par le système SoundRecover de Phonak. (Source : Phonak)

Avant tout il est important de distinguer les différentes techniques et terminologies utilisées, car les termes transposition et compression fréquentielle ne sont pas interprétés de la même façon par tous les auteurs. La transposition est un terme assez général qui illustre simplement le fait qu'une fréquence (ou une bande de fréquences) A est décalée vers une fréquence (ou une zone de fréquences) B. Cette transposition peut être compressive (la bande de sortie est plus étroite que la bande d'entrée) ou expansive (la bande de sortie est alors plus large que la bande d'entrée). La compression fréquentielle, quant à elle, est une façon parmi d'autres de réaliser ce décalage fréquentiel.

La compression fréquentielle consiste à réduire la bande passante du signal de sortie, par rapport au signal d'entrée. Elle peut être linéaire (aussi connue sous le nom de décalage proportionnel) ou non linéaire (décalage non proportionnel).

Le décalage linéaire de fréquence diminue toutes les composantes fréquentielles par un facteur constant et elle offre l'avantage de préserver l'information spectrale, car les ratios entre les différentes composantes fréquentielles sont inchangés par le traitement. De cette manière, les relations entre les fréquences des formants de la parole restent constantes. Ces indices s'avèrent être des indices particulièrement importants pour la reconnaissance des voyelles dans la parole (Neary, 1989). Cependant la hauteur tonale (le pitch) du signal de parole est diminuée et la parole peut avoir une sonorité qui n'est pas naturelle. Par exemple, une voix féminine peut avoir la sonorité d'une voix d'homme. Jusque là le décalage de fréquence linéaire n'a jamais été implémenté dans des systèmes portables (type contour d'oreille), principalement en raison de la sonorité qui n'est pas naturelle.

Le SoundRecover de Phonak est l'exemple d'un algorithme commercialisé (d'abord utilisé dans le Naïda, puis intégré ensuite dans les autres appareils Phonak) de compression non-linéaire de fréquence. Seule la bande passante au-delà de la fréquence de coupure est comprimée et cette fréquence ne peut jamais être inférieure à 1,5 kHz. Cette technique permet de préserver la sonorité naturelle car les graves et medium ne sont pas modifiés. La compression fréquentielle de Phonak repose sur deux principes. Le premier principe consiste à compresser des bandes de fréquences larges dans les hautes fréquences (inaudibles par le patient du fait de la perte auditive) dans une bande de fréquence moins large. Le second principe consiste à ramener cette zone de fréquences comprimées depuis une zone inaudible vers une zone où le patient perçoit encore les sons. C'est-à-dire que les composantes fréquentielles des bandes comprimées seront toutes diminuées en fréquence, par un même ratio. L'enveloppe temporelle du signal n'étant pas dégradée, le signal de sortie sera simplement perçu plus grave que le signal d'entrée. Enfin, il n'y a pas de chevauchement fréquentiel entre la bande basse fréquence et la bande transposée.

Parmi les autres formes de transposition, on distingue des techniques qui consistent à soustraire « n » Hz à toutes les composantes fréquentielles du signal d'entrée (Velmans, 1974 ; Velmans & Marcuson, 1983 ; Rees & Velmans, 1993). Dans certains cas le signal

transposé chevauche le signal basse-fréquence non traité (Robinson et al., 2007 ; 2009). Dans d'autres cas, il n'y a pas de chevauchement fréquentiel entre le signal basse-fréquence non traité et le signal transposé (Simpson et al. 2005 ; 2006).

Actuellement, les derniers appareils qui intègrent de la transposition fréquentielle avec chevauchement sont les Widex et leur algorithme s'appelle Audibility Extender (Kuk, 2007). L'audiogramme du patient est utilisé pour déterminer la fréquence initiale de la transposition. La première fréquence au-delà de 1,6 kHz où le seuil est supérieur à 70 dB HL et qui a une pente supérieure à 10 dB/octave (pour les fréquences entre 0,5 et 4 kHz) est sélectionnée comme la fréquence initiale. Cette fréquence initiale peut également être sélectionnée manuellement par le clinicien entre 0,63 et 6 kHz par intervalle d'1/3 d'octave. Comme pour les méthodes décrites par Velmans (1974) et Robinson et al. (2007), les fréquences inférieures à la fréquence initiale sont amplifiées mais non transposées. Les plus hautes fréquences jusqu'à deux octaves au-dessus de la fréquence initiale sont analysées par l'aide auditive. Une gamme étroite de fréquences ayant les plus hauts niveaux d'intensité dans cette gamme sont sélectionnés, transposés d'une octave, et chevauchent le signal de plus basse fréquence, en dessous de la fréquence initiale. Le clinicien peut ensuite ajuster le niveau du signal transposé et sélectionner une gamme de transposition d'une ou deux octaves au-delà de la fréquence initiale.

L'information basse fréquence n'étant généralement pas affectée par ces méthodes de transposition, ce traitement offre une qualité de son plus naturelle. De plus, le rapport entre les composantes fréquentielles dans les régions transposées est préservé. Le principal défaut de la transposition provient du chevauchement entre l'information haute- et basse-fréquence qui peut nuire à la compréhension, dans la mesure où certaines informations haute-fréquence peuvent masquer l'information basse-fréquence plus utile. De nombreux systèmes essayent désormais d'éviter ce phénomène, en n'autorisant la transposition seulement si le signal d'entrée est composé en majorité d'information haute fréquence.

Elimination de l'effet Larsen

Nous trouvons de plus en plus d'articles et de brevets sur l'intérêt d'utiliser le décalage fréquentiel pour limiter le risque de larsen acoustique (Joson et al., 1993 ; Freed & Soli, 2005,

2009 ; McDermott et al., 2010). Dans la technique proposée par Joson et al., le décalage fréquentiel ne se fait que vers les basses fréquences et il est de l'ordre de 6%. Au contraire, dans les techniques décrites par Freed et Soli et par McDermott, le décalage se fait alternativement vers les basses et hautes fréquences (alternances de phases de compression et d'expansion du signal). Pour Freed et Soli, ce décalage a une valeur de 0,3% de la fréquence. Pour McDermott, il est de l'ordre de 5 Hz. Depuis peu, les appareils auditifs Siemens utilisent la technologie FeedbackStopper qui consiste à appliquer en sortie un décalage fréquentiel de 25 Hz lorsqu'un larsen est détecté. Du fait de ce traitement, il est possible que le patient perçoive une légère variation de sonorité. Cependant, la déficience auditive qui se manifeste en particulier par la perte des cellules ciliées externes, a pour conséquence un élargissement des filtres auditifs. Classiquement, pour une perte auditive moyenne (41 – 70 dB), la largeur des filtres auditifs est 2 fois supérieure à la normale. Pour une perte sévère (71 – 90 dB), la largeur de bande équivalente est environ 4 fois supérieur à la normale (Moore, 2007). On peut alors supposer que ce décalage de 25 Hz dans la zone des 4000 Hz ne sera pas perturbant pour le patient ($df/f = 25/4000 = 6\%$, le seuil de discrimination fréquentiel du normo-entendant étant de l'ordre de 3% au-delà de 1000 Hz).

Chronologie des principaux systèmes et des différentes études

Depuis le début des années 60's, la littérature rapporte divers travaux (articles, brevets) sur les techniques de réduction fréquentielle. La plupart des algorithmes ont été évalués sur diverses populations : chez le malentendant, chez le normo-entendant, ou encore chez le normo-entendant en simulant la perte auditive par filtrage passe bas ou par ajout d'un bruit masquant. Certains des algorithmes ont pu être intégrés dans des systèmes embarqués ou dans des contours d'oreille tandis que d'autres n'ont pas dépassé le stade de l'évaluation en laboratoire, à cause du manque de miniaturisation des appareils. Les résultats de ces études ne montrent en général aucune amélioration significative des performances avec ces systèmes et dans certains cas un simple filtrage passe-bas est tout aussi efficace qu'un algorithme très sophistiqué. Malheureusement la plupart de ces travaux ont été réalisés sur un faible nombre de patient, ce qui ne permet pas toujours de valider statistiquement l'avantage d'un système par rapport à un autre. Dans la suite du présent document, nous présentons différentes techniques imaginées au fil des années dans les différents pays, ainsi que les résultats des principaux articles publiés sur le sujet (cf tableau I en annexe).

Les premiers systèmes de réduction fréquentielle ont été développés en Suède, à Stockholm avec les travaux de Bertil Johansson et Arne Risberg. La prothèse Oticon TP72 (Johansson, 1961) était une des premières à intégrer un traitement de réduction fréquentielle. L'appareil était composé de deux canaux. Les fréquences de 150 Hz à 3 kHz étaient amplifiées normalement. Les fréquences plus hautes entre 4 et 8 kHz étaient passées dans un modulateur non linéaire et converties en bruit basse fréquence en dessous de 1,5 kHz. Le défaut majeur de ce système venait du traitement qui ne permettait pas de préserver les caractéristiques importantes de la forme spectrale des sons d'entrée. Les premiers résultats obtenus avec cet appareil sont présentés dans les études de Johansson (1966) et de Foust et Gengel en (1973) où le transposeur de Johansson a été intégré dans une aide auditive type contour d'oreille. La figure 3 montre des exemples de résultats obtenus dans l'étude de Foust et Gengel, chez 3 sujets, en condition stimulation auditive seule. Le premier exemple (cadre de gauche) montre un sujet qui comprend toujours mieux avec le transposeur qu'avec l'amplification conventionnelle. Le deuxième sujet (cadre du milieu) n'aime pas l'amplification du transposeur et comprend mieux avec son appareil conventionnel qu'avec le transposeur. Enfin le dernier sujet (cadre de droite) a des performances très variables au cours des différentes sessions de test. De ces observations, les auteurs en déduisent que le bénéfice potentiel du transposeur doit être **évalué individuellement**.

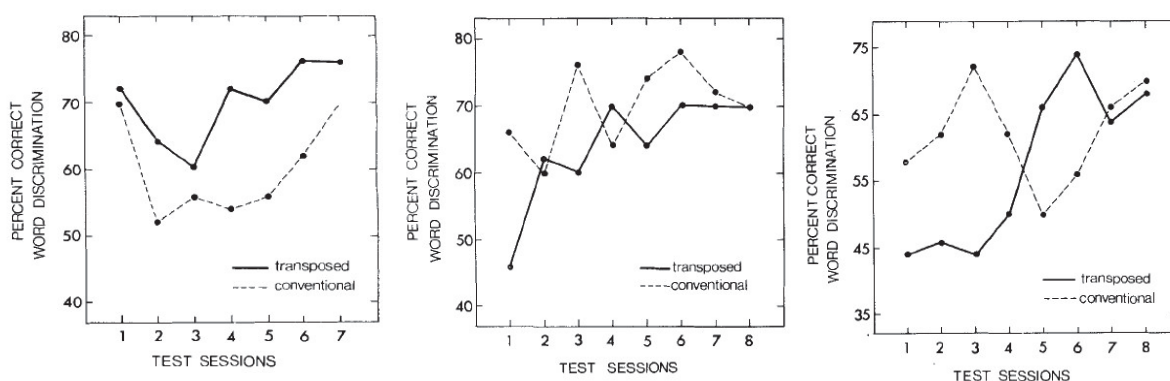


Figure 3. A gauche : exemple d'un sujet qui comprend mieux avec le transposeur qu'avec son appareil conventionnel. Au milieu : exemple d'un sujet qui comprend mieux avec son appareil conventionnel qu'avec le transposeur. A droite : exemple d'un sujet qui a des performances variables. D'après Foust et Gengel (1973).

Quelques années plus tard Risberg (1965) a développé une variante du transposeur [1]. Avec son traitement, les signaux de parole subissaient une transposition puis un filtrage passe-bas à 1000 Hz. En présence de sons basse fréquence type voisement, le transposeur s'éteignait. Pour mesurer l'efficacité de ce système, deux groupes de normo-entendants écoutaient soit des signaux de parole traités par l'algorithme de Risberg, soit des signaux filtrés passe-bas à 500, 750, 1000, 1500, et 2000 Hz. Les auteurs ont mesuré les performances en termes d'intelligibilité des mots et de perception du mode et du lieu d'articulation. Le transposeur n'améliore pas la perception du lieu d'articulation, mais il fournit **100% d'identification correcte pour le mode d'articulation**, ce qui se manifeste en un meilleur score pour la perception des mots (figure 4).

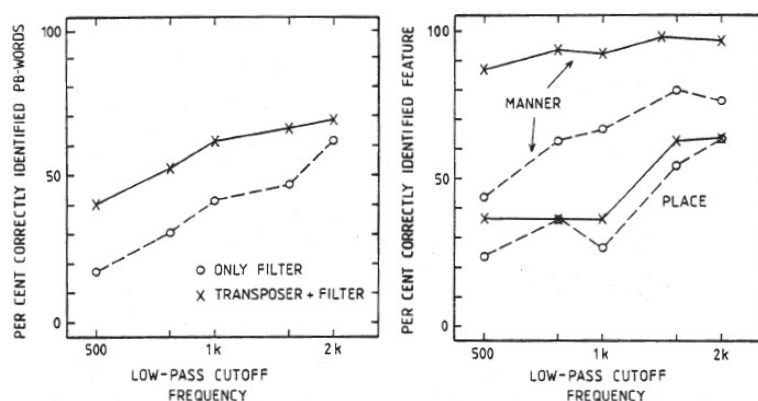


Figure 4. Score de perception des mots et du mode et du lieu d'articulation des syllabes pour un système à transposition fréquentielle versus du filtrage passe-bas. D'après Risberg (1965).

La décennie suivante, en Angleterre, les travaux de Velmans ont conduit au développement du système à transposition FRED, abréviation de Frequency REcoding Device (Velmans, 1974 ; Velmans et Marcuson, 1983). L'appareil divisait le signal entrant en deux bandes, à l'aide d'un passe-bas et un passe-haut, avec une fréquence de croisement à 4 kHz. Une valeur constante de 4 kHz était soustraite de chaque fréquence présente dans la bande passe-haut et les signaux résultants étaient mixés à ceux obtenus par la bande passe-bas, l'information haute-fréquence du spectre n'était donc pas préservée. Ce prototype a fait l'objet de plusieurs études et il ressort de ces travaux que si le transposeur FRED venait à être commercialisé, il deviendra nécessaire de mieux cibler les indications avant de prescrire l'appareil.

Au Japon, Sekimoto et al. (1980) présentent à leur tour une nouvelle variante pour la transposition, il s'agit cette fois d'un système à compression fréquentielle linéaire (figure 5). Les premiers tests réalisés chez le normo-entendant ont montré que le traitement n'améliorait pas l'intelligibilité, la technique a alors été modifiée en une compression non linéaire de fréquence qui réduit la bande passante d'un signal de parole en appliquant des quantités croissantes de diminution fréquentielle à des fréquences d'entrée relativement hautes (Sekimoto et Saito, 1980). Les avantages d'une telle méthode sont 1/ l'absence de chevauchement entre signaux décalés (shiftés) et non décalés, et 2/ une information de basse et moyenne fréquence qui est préservée puisque l'intégralité de la gamme fréquentielle du premier formant est inchangée par le traitement appliqué. Toutefois, la non-linéarité de la compression implique que les ratios de fréquences (incluant les ratios harmoniques) pour ces hautes fréquences qui ont été compressées ne sont plus préservés. La perception de la parole peut, ainsi, être négativement affectée si la fréquence de coupure de la compression était diminuée au point d'inclure des plus basses fréquences.

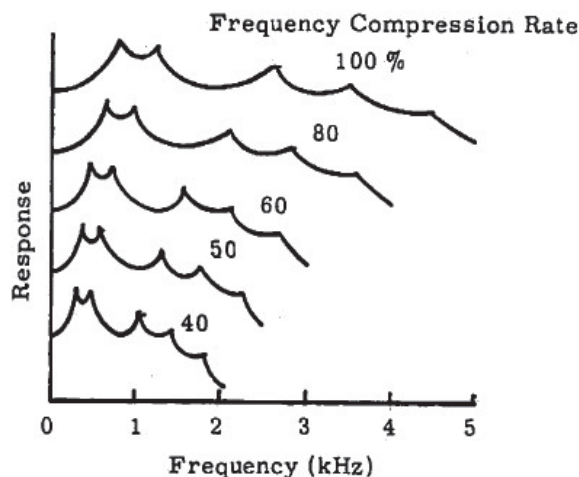


Figure 5. Enveloppes d'un spectre fréquentiel de parole pour différents taux de compression fréquentielle linéaire. D'après Sekimoto et al. (1980b).

Parallèlement, en France on trouve le système Galaxie qui a été pensé par Jean-Claude Lafon. Depuis le milieu des années 60's Lafon a déposé plusieurs brevets (avec la collaboration de Gharbi pour certains) en utilisant divers systèmes de compression fréquentielle linéaire et non linéaire. Le but des traitements utilisés est de présenter l'information dans « le spectre audible

par le sourd » (soit de 20 Hz à 1000 Hz) et de préserver la structure prosodique des signaux de parole. Pour cela, différentes bandes spectrales sont transposées dans des zones de plus basses fréquences et amplifiées. Le nombre de bandes de destination est égal au nombre de bandes sources. Le ratio de compression spectrale est différent pour chacune des bandes mais il est constant à l'intérieur d'une même bande. Les fréquences en dessous de 500 Hz sont amplifiées avec un gain constant. L'amplification diminue graduellement entre 500 et 1000 Hz, de façon à libérer cette zone pour les bandes transposées. Plusieurs prototypes ont été conçus, mais aucun n'a été commercialisé à cause du manque de miniaturisation de ces dispositifs. D'après les auteurs, ces systèmes fourniraient un bénéfice mais il n'a jamais vraiment été quantifié.

Un peu plus tard, toujours en France, Dupret et Lefevre (1991) déposent le brevet de leur prothèse EMILY. Le principe de celle-ci est de déterminer le second formant F2 par extraction des fréquences dominantes dans la bande [1 kHz – 2 kHz]. F2 est ensuite additionné avec $F2/2$ et $F2*2$ au signal original, et chaque composante F2 peut être amplifiée individuellement en fonction de l'audiogramme du patient. L'algorithme était implémenté dans un système embarqué et il était utilisé avec des appareils Phonak. D'après ses inventeurs, le système était sensé améliorer l'intelligibilité de la parole en milieu calme et dans le bruit, mais aucun article n'a jamais été publié sur l'utilité de ce système.

Aux Etats-Unis, à Cambridge, Reed et al. (1983, 1985) développent également des algorithmes qui réalisent une réduction fréquentielle linéaire ou non-linéaire [2], mais cette technique ne dépassera jamais le stade de l'évaluation en laboratoire. Dans les algorithmes utilisés, la fréquence fondamentale est préservée car l'espacement entre les périodes successives du pitch sont maintenues. La bande passante des signaux comprimés est soit 2500 Hz, soit 1250 Hz. Malheureusement les résultats pour les différentes conditions expérimentales (figure 6) montrent qu'un simple filtrage passe-bas (à bande-passante équivalente) est tout aussi efficace, voire plus efficace que leur algorithme.

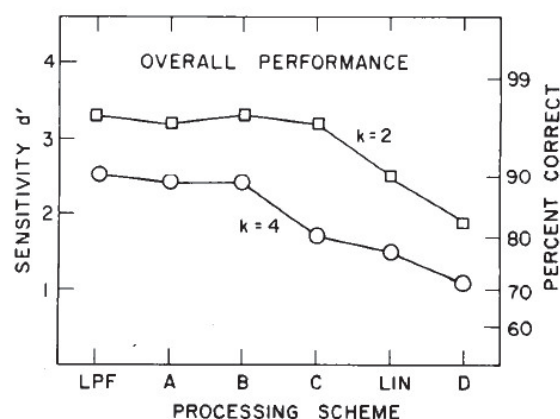


Figure 6. Performances moyennes obtenues sur 6 sujets, pour 6 conditions de réduction fréquentielle a et 2 facteurs de compression K . LPF correspond au filtrage passe-bas à 2500 Hz pour $K = 2$ et 1250 Hz pour $K = 4$. Les labels A, B, C, LIN et D correspondent à des valeurs du paramètre $a = 0,5 ; 0,33 ; 0,16 ; 0 ;$ et $-0,33$ pour $K = 2$ et à $a = 0,7 ; 0,6 ; 0,3 ; 0 ;$ et $-0,3$ pour $K = 4$. D'après Reed et al. (1983).

Vers le milieu des années 90, la société AVR Communications propose plusieurs appareils reposant sur différentes techniques de transposition fréquentielle. Il y a d'abord le système Transonic apparu en 1997. Ce processeur décale proportionnellement et comprime la fréquence du signal de parole dans une gamme de fréquences plus basse. Le processeur détermine si les pics spectraux d'énergie sont localisés au-dessus ou en dessous de 2500 Hz. Si l'énergie spectrale est au-delà de 2500 Hz, le signal est considéré comme une consonne et le coefficient de consonne (Z_c) est activé. Si l'énergie spectrale est en dessous de 2500 Hz, le signal est considéré comme une voyelle et le coefficient de voyelle (Z_v) est activé. L'activation du coefficient Z_c ou Z_v se manifeste en un décalage du spectre de la parole dans une gamme de fréquences plus basses. En complément, un circuit DCB (Dynamic Consonant Boost) est ajusté pour augmenter le niveau des consonnes les plus faibles. Toutefois les résultats présentés par Parent et al. (1997) pour la compréhension avec ce système sont plutôt mitigés.

Quelques années plus tard, en 2000, AVR sort la prothèse ImpaCt. A la différence du Transonic, cet appareil ne détecte pas les consonnes et voyelles mais il détecte les sons voisés et non voisés. La transposition n'est active que si l'appareil détecte un son non-voisé en entrée de l'appareil et la quantité par laquelle les fréquences sont divisées quand la transposition est active est appelée FCVL (Frequency Compression for Voiceless Sounds). Par exemple,

lorsque FCVL est fixé à 2 et la transposition est active, toutes les fréquences d'entrée sont décalées d'une octave vers les graves.

Enfin, l'appareil AVR le plus récent semble être le Sonovation Nano Xp, distribué depuis 2004. Le principe de ce dernier système n'est pas détaillé dans la littérature, mais d'après Gifford et al. (2007), il semble reposer sur la même philosophie que l'ImpaCt.

En Australie, depuis 2005 l'équipe de McDermott et de Simpson utilisent un nouveau processeur expérimental de compression fréquentielle non linéaire (le SHARP, Zakis et al. 2001) qui vise à laisser intacte l'information basse fréquence, inférieure à la fréquence de coupure de l'appareil. C'est d'ailleurs cet algorithme qui a donné naissance au SoundRecover de Phonak. L'équipe de Simpson a réalisé des tests sur une vingtaine de malentendants et les résultats montrent globalement de meilleures performances avec cette technique (Simpson et al., 2005 ; 2006). Cependant les auteurs décrivent une très large variabilité des résultats et précisent qu'il est nécessaire de bien maîtriser les facteurs de compression fréquentielle pour obtenir un réglage optimal.

Parallèlement Widex commence à proposer son algorithme Audibility Extender. Les principaux auteurs à travailler avec ce système (intégré dans les appareils Inteo et Mind) sont Francis Kuk et ses collaborateurs (Kuk et al, 2006, 2007, 2009). Comme nous l'avons décrit dans l'introduction, la transposition de Widex consiste en une transposition linéaire dont la fréquence de coupure, réglable par le clinicien, se situe entre 630 et 6000 Hz. La bande fréquentielle transposée est décalée d'une voire deux octaves dans les plus basses fréquences et chevauche la bande fréquentielle non traitée. Kuk et al notent qu'1/3 des sujets testés trouvent l'écoute plus confortable lorsque la transposition est activée. Dans leur étude de 2009 où ils évaluent la compréhension de la parole dans le silence et dans le bruit, on constate que les résultats s'améliorent au bout de 2 mois d'utilisation des appareils (d'où la nécessité d'un entraînement), mais les analyses statistiques montrent que seule la perception des fricatives est améliorée par la transposition.

Récemment en Inde, Kulkarni et al. (2007 et 2009) ont proposé une technique de compression fréquentielle multi-bande, en vue de réduire les effets du masquage fréquentiel [4]. Usuellement, les appareils à compression non linéaire de fréquence travaillent sur deux bandes, une bande de basse fréquence jusqu'à la coupure de la compression qui reste intacte et une seconde bande de plus haute fréquence qui subit l'opération de compression fréquentielle. A la différence de ces techniques, l'algorithme proposé par Kulkarni et al consiste à diviser le signal d'entrée de l'appareil en plusieurs bandes et à réaliser la compression fréquentielle à l'intérieur des différentes bandes. Cette compression multi-bande concentre ainsi l'énergie spectrale vers les centres des bandes pour compenser partiellement le masquage spectral. Les résultats de ce système testé chez le normo-entendant montrent un avantage de la compression seulement en condition bruyante où le bruit est plus fort que la parole (figure 7, partie gauche). Chez le malentendant (figure 7, partie droite), la compression est généralement meilleure que l'amplification conventionnelle et les premiers résultats sont plutôt encourageants.

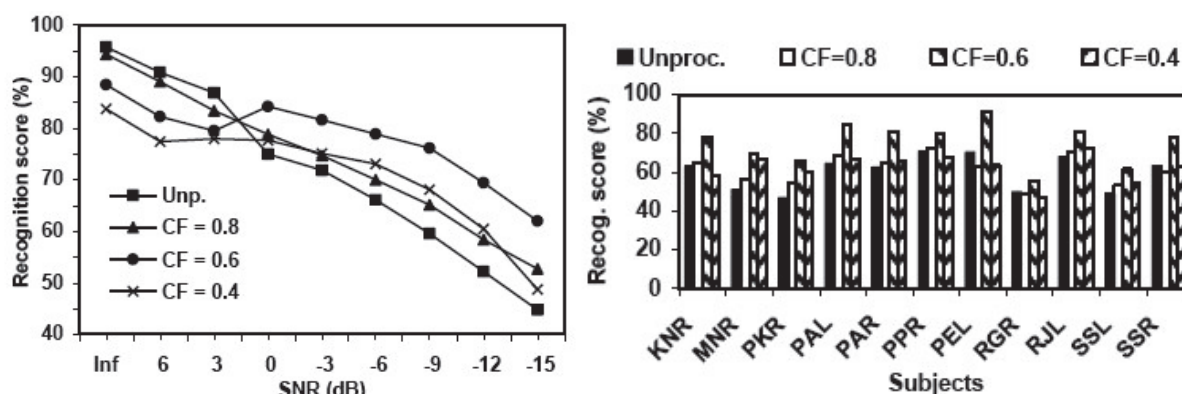


Figure 7. Intelligibilité des mots CNC pour différents facteurs de compression (1, 0,8, 0,6 et 0,4). A gauche : scores moyens en fonction du RSB mesurés chez 6 sujets normo-entendants. A droite : scores individuels mesurés dans le calme chez 11 sujets malentendants.

A cette même période, l'équipe de Brian Moore au Royaume-Uni s'interroge sur l'intérêt d'utiliser la transposition fréquentielle pour l'appareillage des sourds partiels qui ont des ZICs dans les hautes fréquences. Robinson et al entreprennent alors des travaux de recherche pour évaluer les performances offertes par un nouvel algorithme expérimental de transposition fréquentielle [5], d'abord en laboratoire (2007), puis embarqué dans des aides auditives

Phonak Aero 311 (2009). Pour cet algorithme, le choix de la fréquence de coupure de leur transposition (la fréquence en dessous de laquelle le signal est inchangé) est basé sur un critère recommandé par Vickers et al. (2001) et par Baer et al. (2002) qui recommandent de ne pas amplifier les fréquences au-delà de $1,7*f_e$, f_e étant la fréquence charnière de la ZIC de chaque patient. Les patients sont donc testés avec une condition contrôle pour laquelle l'aide auditive amplifie les fréquences jusqu'à $1,7*f_e$ (c'est-à-dire type filtrage passe-bas), et une condition transposée pour laquelle le signal entre $[2*f_e - 2,7*f_e]$ est transposé-ajouté sur la zone $[f_e - 1,7*f_e]$ (figure 8). La zone transposée n'est pas comprimée, elle garde une largeur de bande de $0,7*f_e$. La transposition ne s'active que lorsque le ratio [énergie haute fréquence / énergie basse fréquence] devient supérieur à un certain seuil (0,1 dans Robinson et al, 2007 ; 0,5 dans Robinson et al, 2009). Tout comme l'étude de Reed et al. (1993), les résultats chez les malentendants révéleront qu'un simple filtrage passe-bas est tout aussi efficace que l'algorithme étudié. La transposition améliore la perception de certaines consonnes, mais cela se fait au détriment de la perception des voyelles.

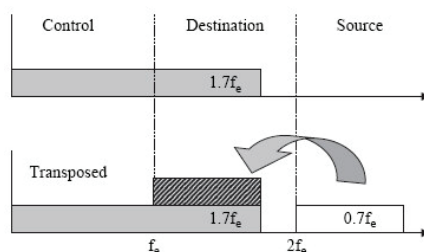


Figure 8. Illustration du traitement réalisé par l'Algorithme de Robinson et al. (2007, 2009). La condition contrôle est un filtrage passe-bas à $1,7*f_e$. Pour la condition transposée, la bande source entre $[2*f_e - 2,7*f_e]$ est transposée sur la zone $[f_e - 1,7*f_e]$. D'après Robinson et al. (2007).

En Chine on trouve les travaux de Xiao et al. (2008) qui ont comparé les performances de deux algorithmes par rapport à une amplification conventionnelle. L'algorithme M1 consiste en une compression fréquentielle couplée à un chevauchement du signal transposé sur la partie haute fréquence du signal non transposé. L'algorithme M2 consiste à compresser séparer les bandes basse et haute fréquence et à placer le contenu haute fréquence compressé derrière l'information compressée basse fréquence, d'où une absence de chevauchement fréquentiel entre les bandes. Six sujets malentendants avec une perte modérée à sévère étaient recrutés pour évaluer ces deux algorithmes. Les résultats montrent que les patients ont en

général de meilleurs scores avec les algorithmes M1 et M2 qu'avec leur propre appareil auditif. Les auteurs notent qu'en milieu bruyant, les systèmes à compression sans chevauchement fréquentiel sont plus stables que les systèmes à transposition avec un chevauchement fréquentiel.

Au Canada, Glista et al. (2009) évaluent un nouveau traitement de compression non linéaire de fréquence implémenté dans des contours d'oreille type Phonak Savia 311 ou 411, chez 24 sujets ayant une surdité sévère à profonde dans les hautes fréquences. Le fonctionnement de l'algorithme n'est pas détaillé mais les auteurs précisent simplement que la fréquence de coupure et le ratio de compression sont ajustés à chaque sujet. Les résultats moyens pour la tâche de détection des sons /s/ et /sh/ sont présentés figure 9, partie gauche. La partie droite de la figure X affiche les résultats individuels pour la détection des sons, tracé en terme de différence de scores. Les scores négatifs indiquent une amélioration avec le traitement conventionnel (CP) et les scores positifs indiquent une amélioration avec la NFC. Cette figure indique une très large variabilité du bénéfice fourni par le traitement conventionnel et par la compression fréquentiel, chez les participants adultes et enfants. Les résultats aux tâches de reconnaissance de parole (consonnes, pluriels, voyelles) montrent que la transposition conduit à une amélioration de la perception des consonnes et surtout elle ne nuit pas à la perception des voyelles (figure 10), comme c'était le cas dans l'étude de Robinson et al. (2007).

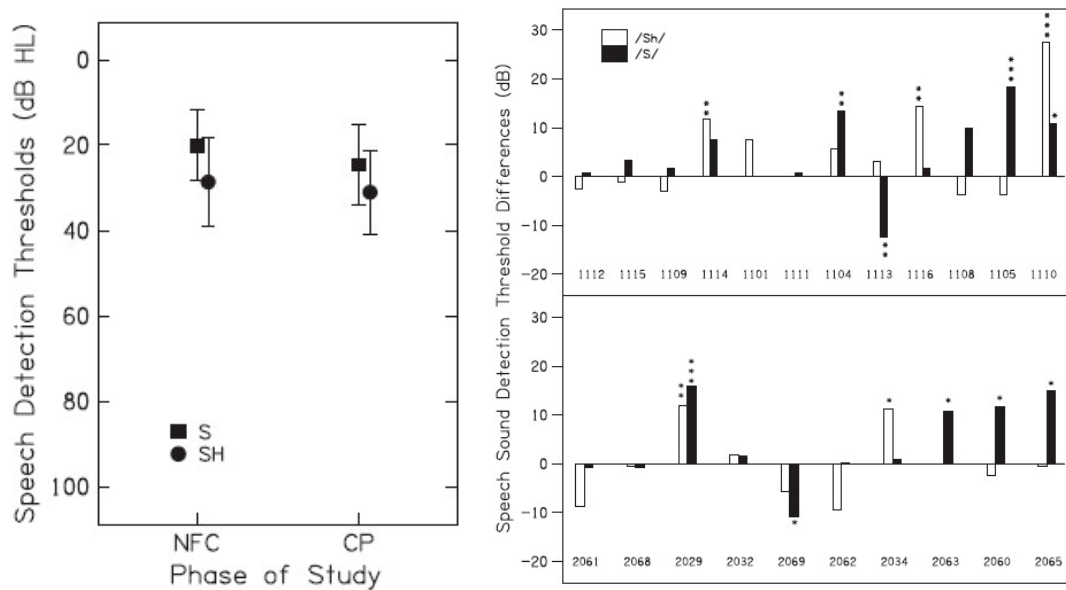


Figure 9. Partie gauche : seuil moyen de détection des sons /s/ et /sh/ chez les 24 participants (13 adultes et 11 enfants), pour le traitement conventionnel (CP) et pour la compression non linéaire de fréquence (NFC). Partie Droite : seuils individuels de détection des sons /s/ et /sh/ exprimés en terme de différence de score. Résultats chez les adultes (en haut) et chez les enfants (en bas). Un score négatif indique une amélioration avec le CP, un score positif indique une amélioration avec la NFC. D'après Glista et al. (2009).

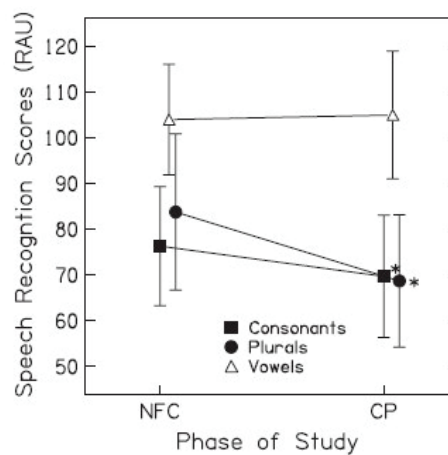


Figure 10. Scores moyens de reconnaissance de la parole chez les 24 participants (13 adultes et 11 enfants), pour le traitement conventionnel (CP) et pour la compression non linéaire de fréquence (NFC). D'après Glista et al. (2009).

Enfin, une des dernières études en date a été réalisée par les équipes Allemandes et Suisses de Bohnert et al. (2010) pour l'évaluation du SoundRecover de Phonak dont nous avons décrit le fonctionnement en introduction. Onze sujets malentendants ayant une perte sévère à profonde dans les hautes fréquences étaient testés avec un appareil Naïda. Les tests consistaient à reconnaître les phrases OLSA dans un bruit qui avait le même spectre à long terme que le matériel phonétique utilisé comme stimulus. Sur les 11 sujets testés, 7 comprennent mieux avec le SoundRecover et 4 comprennent moins bien (figure 11), la dégradation du seuil d'intelligibilité pouvant atteindre 7.85 dB chez le sujet S2. Cette étude montre une fois de plus que tous les patients ne tirent pas un bénéfice de la compression fréquentielle.

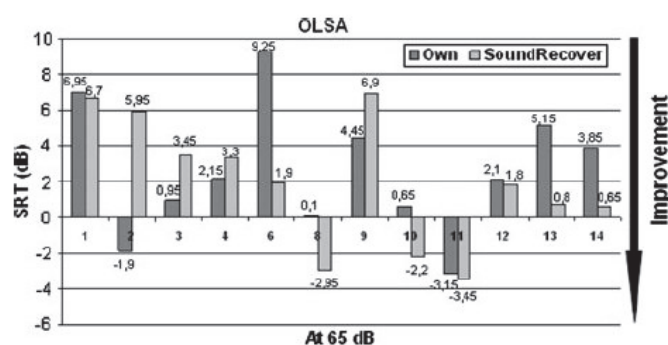


Figure 11. Seuil de reconnaissance de la parole (SRT). Les barres en gris foncé montrent les résultats des sujets avec leurs propres aides auditives. Les barres en gris clair montrent les résultats obtenus avec le SoundRecover. D'après Bohnert et al. (2010).

L'ensemble des données de la littérature sur la réduction fréquentielle décrite ci-avant et sur le tableau I montrent une très large disparité des performances, quelque soit la technique utilisée et quelque soit le degré de perte auditive dans les hautes fréquences. Comment expliquer alors les succès et les échecs de ces techniques ? D'après les différents auteurs, certains échecs peuvent s'expliquer par les distorsions du son engendrées par les traitements réalisés par ces systèmes. Par exemple Robinson et al. (2007) expliquent l'échec de l'ImpaCt à cause du traitement qu'il réalise. La technique de « slow playback » va diminuer toutes les fréquences de sortie, ce qui résulte en un allongement du signal temporel par rapport à l'entrée. Pour éviter une désynchronisation de la sortie par rapport à l'entrée, certains segments doivent être régulièrement supprimés, ce qui a pour conséquence des distorsions du signal de sortie par rapport à l'original et des pertes d'information.

D'autres auteurs supposent que les performances sont liées aux paramètres de la compression et qu'il est nécessaire de respecter certaines règles pour ne pas dégrader la compréhension. Dans les études de Kulkarni et al. (2007, 2009) par exemple, nous constatons qu'une compression fréquentielle trop importante (0,4) peut nuire à la compréhension par rapport à un ratio optimal de 0,6. Il est donc indispensable de respecter certaines règles et de savoir maîtriser les différents paramètres de la compression fréquentielle car les performances au niveau perceptif peuvent être sévèrement affectées (Baskent & Shannon, 2007).

On trouve dans la littérature plusieurs recommandations cliniques. Selon Turner & Hurtig (1999) la compression fréquentielle doit être utilisée lorsque la perte devient supérieure à 60 dB HL au-delà de 2000 Hz. Ils préconisent que l'information haute fréquence de la parole soit décalée vers les régions de plus basses fréquences, où la sensibilité auditive est meilleure que 60 dB HL. Quant à l'optimisation du réglage avec la compression, Turner et Hurtig recommandent que :

- 1) le décalage de fréquences se fasse avec un facteur multiplicatif fixe,
- 2) les caractéristiques temporelles (enveloppe) et la durée de la parole soient préservées autant que possible,
- 3) la bande de fréquence comprimée ne soit pas inférieure à 60-70% de la largeur de bande originale, c'est-à-dire que le facteur de compression ne doit pas être supérieur à 1,4 ($1 / 0,7 = 1,4$), sans quoi l'intelligibilité du signal risque d'être dégradée.

Robinson et al. (2007, 2009) suivent les recommandations de Vickers et al. (2001) et de Baer et al. (2002) et n'amplifient pas les fréquences supérieures à $1,7 \cdot f_e$. Enfin, si on s'en tient aux résultats de Xiao et al. (2008), on peut penser qu'en milieu bruyant, il est préférable d'utiliser des systèmes à compression sans chevauchement fréquentiel.

Depuis une cinquantaine d'années, les systèmes à transpositions fréquentielles se sont succédés. La plupart se sont révélés inefficaces, mais ces échecs étaient probablement dus aux technologies qui étaient inadaptées à l'époque.

La technologie est désormais stabilisée et les systèmes disponibles aujourd'hui sont très prometteurs. Ils offrent un bénéfice majeur pour la gestion du larsen et pour l'audibilité des hautes fréquences qui n'est pas nécessairement reliée à l'intelligibilité (entendre les bruits

environnants, les oiseaux, les cigales, le clignotant). Ils semblent être très prometteurs pour les surdités sévères et profondes et sont sans conteste des solutions d'avenir pour les surdités de type pente de ski.

Aujourd'hui encore il existe de larges variabilités de performances (Bohnert et al., 2010) et on peut espérer que la recherche, ces prochaines années, s'orientera sur la sélection des candidats à la transposition. Les indications d'appareillage sont encore trop peu connues. Ce point crucial sera peut-être le dernier obstacle à surmonter pour obtenir un résultat optimal avec les nouvelles techniques de réduction fréquentielle.

Références :

- Baer T, Moore BC, Kluk K. Effects of low pass filtering on the intelligibility of speech in noise for people with and without dead regions at high frequencies. *J Acoust Soc Am*. 2002;112:1133-44.
- Başkent D, Shannon RV. Combined effects of frequency compression-expansion and shift on speech recognition. *Ear Hear*. 2007;28:277-89.
- Bohnert A, Nyffeler M, Keilmann A. Advantages of a non-linear frequency compression algorithm in noise. *Eur Arch Otorhinolaryngol*. 2010;267:1045-53.
- Ching TY, Dillon H, Byrne D. Speech recognition of hearing-impaired listeners: predictions from audibility and the limited role of high-frequency amplification. *J Acoust Soc Am*. 1998;103:1128-40.
- Ching TY, Dillon H, Katsch R, Byrne D. Maximizing effective audibility in hearing aid fitting. *Ear Hear*. 2001;22:212-24.
- Dupret JP, Lefevre F. Electronic device for processing a sound signal. US patent US 5077800, 1991.
- Foust KO, Gengel RW. Speech discrimination by sensorineural hearing-impaired persons using a transposer hearing aid. *Scand Audiol*. 1973;2:161-70.
- Freed DJ, Soli SD. Frequency shifter for use in adaptive feedback cancellers for hearing aids. US patent US 2005/0271222 A1.
- Freed DJ, Soli SD. Frequency shifter for use in adaptive feedback cancellers for hearing aids. US patent US 7609841 B2, 2009.
- Gifford RH, Dorman MF, Spahr AJ, McKarns SA. Effect of digital frequency compression (DFC) on speech recognition in candidates for combined electric and acoustic stimulation (EAS). *J Speech Lang Hear Res*. 2007;50:1194-202.
- Glista D, Scollie S, Bagatto M, Seewald R, Parsa V, Johnson A. Evaluation of nonlinear frequency compression: clinical outcomes. *Int J Audiol*. 2009;48:632-44.
- Hogan CA, Turner CW. High-frequency audibility: benefits for hearing-impaired listeners. *J Acoust Soc Am*. 1998;104:432-41.
- Johansson B. A new coding amplifier system for the severely hard of hearing. 1961. Proceedings of the 3rd International Congress on Acoustics, Stuttgart 1959;2:655-657.
- Johansson B. The use of the transposer for the management of the deaf child. *Int J Audiol*. 1966;5:362-72.

- Joson HA, Asano F, Suzuki Y, Sone T. Adaptive feedback cancellation with frequency compression for hearing aids. *J Acoust Soc Am*. 1993;94:3254-8.
- Kuk F, Keenan D, Korhonen P, Lau CC. Efficacy of linear frequency transposition on consonant identification in quiet and in noise. *J Am Acad Audiol*. 2009;20:465-79.
- Kuk F, Korhonen P, Peeters H, Keenan D, Jessen A, Andersen H. Linear frequency transposition: Extending the audibility of high frequency information. *Hear Rev*. 2006;13:42-8.
- Kuk F, Peeters H, Keenan D, Lau C. Use of frequency transposition in a thin-tube open-ear fitting. *Hearing Journal* 2007;60:59-63.
- Kulkarni PN, Pandey PC, Jangamashetti DS. Multi-band frequency compression for reducing the effects of spectral masking. *Int J Speech Technol*. 2007;10:219–27.
- Kulkarni PN, Pandey PC, Jangamashetti DS. Multi-band frequency compression for sensorineural hearing impairment. *Proc. 16th IEEE International Conference on Digital Signal Processing 2009 (DSP 2009)*, Santorini, Greece. 2009 :1-6.
- Lafon JC. Hearing aids. US patent US 3385937, 1968.
- Lafon JC. Perfectionnement aux dispositifs de prothèse auditive. EU patent EP 0054450 B1, 1984.
- Lafon JC, Gharbi T. Perfectionnement aux dispositifs de prothèse auditive. EU patent EP 0246970 B1, 1987.
- Lafon JC, Gharbi T. Hearing aid devices in which high frequency signal portions are transposed in low frequency compension signal portions. US patent US 4843623, 1989.
- McDermott HJ, Dorkos VP, Dean MR, Ching TY. Improvements in speech perception with use of the AVR TranSonic frequency-transposing hearing aid. *J Speech Lang Hear Res*. 1999;42:1323-35.
- McDermott H, Hersbach A, Simpson A. Feedback suppression in sound signal processing using frequency translation. US patent US 7778426 B2, 2010.
- McDermott HJ, Knight MR. Preliminary results with the AVR ImpaCt frequency-transposing hearing aid. *J Am Acad Audiol*. 2001;12:121-7.
- Moore BCJ. Cochlear hearing loss - Physiological, Psychological and Technical Issues. Second edition, Wiley, 2007.
- Moore BCJ. Dead Regions in the Cochlea: Diagnosis, Perceptual Consequences, and Implications for the Fitting of Hearing Aids. *Trends Amplif*. 2001;5:1-34.
- Moore BC, Huss M, Vickers DA, Glasberg BR, Alcántara JI. A test for the diagnosis of dead regions in the cochlea. *Br J Audiol*. 2000;34:205-24.
- Nearey TM. Static, dynamic, and relational properties in vowel perception. *J Acoust Soc Am*. 1989;85:2088-113.
- Parent TC, Chmiel R, Jerger J. Comparison of performance with frequency transposition hearing aids and conventional hearing aids. *J Am Acad Audiol*. 1997;8:355-65.

- Reed CM, Hicks BL, Braida LD, Durlach NI. Discrimination of speech processed by low-pass filtering and pitch-invariant frequency lowering. *J Acoust Soc Am*. 1983;74:409-19.
- Reed CM, Schultz KI, Braida LD, Durlach NI. Discrimination and identification of frequency-lowered speech in listeners with high-frequency hearing impairment. *J Acoust Soc Am*. 1985;78:2139-41.
- Rees R, Velmans M. The effect of frequency transposition on the untrained auditory discrimination of congenitally deaf children. *Br J Audiol*. 1993;27:53-60.
- Risberg A. The 'transposer' and a model of speech perception. *Quarterly Progress and Status Report STL-QPSR* 1965 ;4 :26-30.
- Robinson JD, Baer T, Moore BC. Using transposition to improve consonant discrimination and detection for listeners with severe high-frequency hearing loss. *Int J Audiol*. 2007;46:293-308.
- Robinson JD, Stainsby TH, Baer T, Moore BC. Evaluation of a frequency transposition algorithm using wearable hearing aids. *Int J Audiol*. 2009;48:384-93.
- Sakamoto S, Goto K, Tateno M, Kaga K. Frequency compression hearing aid for severe-to-profound hearing impairments. *Auris Nasus Larynx*. 2000;27:327-34.
- Sekimoto S, Kiritani S, Saito S. Intelligibility of frequency compressed speech in lowpass filtered condition. *Ann Bull*. 1980;14:181-93.
- Sekimoto S, Saito S. Nonlinear frequency compression speech processing based on the PARCOR analysis-synthesis technique. *Ann Bull*. 1980;14:65-72.
- Simpson A, Hersbach AA, McDermott HJ. Frequency-compression outcomes in listeners with steeply sloping audiograms. *Int J Audiol*. 2006;45:619-29.
- Simpson A, Hersbach AA, McDermott HJ. Improvements in speech perception with an experimental nonlinear frequency compression hearing device. *Int J Audiol*. 2005;44:281-92.
- Turner CW. Hearing loss and the limits of amplification. *Audiol Neurotol*. 2006;11 Suppl 1:2-5.
- Turner CW, Hurtig RR. Proportional frequency compression of speech for listeners with sensorineural hearing loss. *J Acoust Soc Am*. 1999;106:877-86.
- Velmans M. The design of speech recoding devices for the deaf: University of Manchester. *Br J Audiol*. 1974;8:1-5.
- Velmans M, Marcuson M. The acceptability of spectrum-preserving and spectrum-destroying transposition to severely hearing-impaired listeners. *Br J Audiol*. 1983;17:17-26.
- Vickers DA, Moore BC, Baer T. Effects of low-pass filtering on the intelligibility of speech in quiet for people with and without dead regions at high frequencies. *J Acoust Soc Am*. 2001;110:1164-75.
- Xiao X, Hu G, Liu C, Liu J. Improving speech intelligibility of severely hearingimpaired people by frequency-lowering technique. *Scientific Research and Essay* 2008;3:259-66.

Zakis JA, McDermott HJ, Fisher M. A high performance digital hearing aid for advanced sound processing research. 2001. Poceedings of “Biomedical Research in 2001”, IEEE Engineering in Medicine and Biology Society (Victorian Chapter), Monash University, Australia, 17-20.

Annexe :

Article Auteurs et année	Système utilisé : Traitement custom ou du commerce	Population NE: normo- entendants ME : malentendants	Mesure Matériel	Conditions mesurées : T: transpo activée NT: transpo désactivée/ PB: passe- bas en Hz	Performances moyennes <, >, ou =, individuelles si disponibles	Résultat principal Remarques des auteurs
Risberg, 1965	Personnalisé [1]	NE (2 groupes)	reconnaissance des mots, perception du mode et du lieu d'articulation	T, PB500, PB750, PB1000, PB2000	- lieu d'articulation: T = PB; - mode d'articulation : T > PB; - perception des mots : T > PB	- le transposeur n'améliore pas la perception du lieu d'articulation; - 100% d'identification correcte pour le mode d'articulation; - meilleur score pour la perception des mots
Johansson, 1966	Oticon TP72	enfants ME (une vingtaine, répartis en différents groupes)	perception des mots et des consonnes type fricatives, divers matériels phonétiques	NT, T	Dans la plupart des cas, T > NT mais seulement après plusieurs séances d'entraînement	- meilleure discrimination de /ʃa/ vs. /sa/ en condition transposée; - tous les sujets ont besoin d'un entraînement pour apprendre à comprendre avec le système; - lorsque le transposeur est éteint à la fin de l'expérimentation, tous les sujets voient leurs performances se dégrader; - l'appareil est très sensible au bruit
Foust & Gangel, 1973	Oticon TP72	9 ME avec une surdit� partielle congénitale ou acquise	reconnaissance des mots	NT, T	T = NT, mais les données individuelles sont très variables	- la discrimination du phon�me /s/ est toujours meilleure avec la transposition activ�e; - un entra�nement intensif est n�cessaire pour comprendre la parole transpos�e; - le b�n�fice potentiel du transposeur doit �tre �valu� individuellement

Sekimoto et al., 1980	compression linéaire de fréquence	5 NE avec simulation de perte auditive par des filtrages passe-bas à 700, 1000 et 1500 Hz	perception des consonnes et des voyelles prononcées par des hommes et des femmes	NT, T avec des ratios de compression de 80, 60, 50 et 40%	- consonnes ♂ : =; - consonnes ♀ : =; - voyelles ♂ : =; - voyelles ♀ : +	- amélioration des performances seulement pour la perception des voyelles prononcées par des voix féminines
Velmans & Marcuson, 1983	FRED, Oticon TP72	36 ME	reconnaissance des mots VCV sans signification	T-FRED, T-TP72	- pour le FRED, 34/36 sujets comprennent mieux avec la transposition activée - pour le TP72, 28/36 sujets comprennent mieux sans transposition	- les sujets préfèrent généralement la transposition du FRED à celle du TP72; - le son est trop distordu avec le TP72; - un entraînement intensif est nécessaire pour comprendre avec l'algorithme de transposition
Reed et al., 1983	personnalisé [2], compression linéaire et non-linéaire de fréquence, sans chevauchement fréquentiel	2 groupes de 3 NE	reconnaissance des syllabes CV	PB1250, PB2500, T pour différentes conditions	T = PB	- la transposition est aussi efficace qu'un simple filtrage passe-bas ; - une rééducation auditive assez longue est indispensable pour apprendre à « encoder » (à comprendre) les signaux de parole diminués en fréquence
Reed et al., 1985	personnalisé [2], pas de chevauchement fréquentiel	4 ME avec une surdité dans les hautes-fréquences et 4 NE	reconnaissance des syllabes CV	PB1250, PB2500, T pour différentes conditions	NT > T	- l'analyse des confusions phonétiques montre que la réduction fréquentielle améliore la perception des fricatives et dégrade la perception des semi-voyelles; - la compression fréquentielle n'est pas meilleure que l'amplification linéaire

Rees & Velmans, 1993	FRED	8 ME, jeunes sourds congénitaux	reconnaissance des mots monosyllabiques et des syllabes sans signification. 2 séances test-retest espacées de 1 à 7 jours	T, NT	<ul style="list-style-type: none"> - en général T > NT (+8%); - pour 7 sujets T > NT (+3% jusqu'à +16%); - pour un sujet T < NT (-6%); - test = retest 	<ul style="list-style-type: none"> - pas de différence significative entre les sessions test et re-test --> le FRED améliore significativement les capacités des sujets à discriminer les consonnes, sans pour autant avoir été familiarisé au préalable; - les scores peuvent varier considérablement d'un sujet à l'autre; - la transposition peut parfois conduire à un masquage partiel des indices basse fréquence de la parole par le signal transposé, d'où une baisse de performances observée chez certains patients; - il est nécessaire de mieux cibler les indications avant de prescrire l'appareil
Parent et al., 1997	AVR Transonic	4 ME avec une surdité partielle	audiométrie tonale, reconnaissance des spondees (mots dissyllabiques), des voyelles, des consonnes, des phrases et des mots monosyllabiques	NT, T	<ul style="list-style-type: none"> - pour les sujets S1 et S2 T > NT; - pour le sujet S3 T = NT ; - pour le sujet S4 T < NT 	<ul style="list-style-type: none"> - le Transonic améliore la perception des hautes fréquences, mais les résultats pour la compréhension sont plutôt mitigés --> une meilleure audibilité n'implique pas nécessairement une meilleure intelligibilité
McDermott et al., 1999	AVR Transonic	5 ME avec une surdité partielle (perte > 80 dB au-delà de 1000 Hz)	audiométrie tonale, reconnaissance des phrases (mots clés) et des consonnes	NT, T avec différents paramétrages de la transposition	<ul style="list-style-type: none"> - globalement T > NT; - pour le sujet S5, T < NT pour certains paramétrages 	<ul style="list-style-type: none"> - tous les patients ne tirent pas un bénéfice de la compression

Turner & Hurtig, 1999	compression linéaire de fréquence	15 ME (16 oreilles) + 3 NE	reconnaissance des consonnes prononcées par des hommes et des femmes	NT, T avec des ratios de compression de 0,9; 0,8; 0,7; 0,6; 0,5	- globalement T > NT; - seuls 6/15 sujets ont une amélioration significative avec T (au moins 6% d'écart entre T et NT)	- la compression de fréquence proportionnelle couplée à l'amplification peut offrir des améliorations significatives des résultats par rapport à l'amplification seule
Sakamoto et al., 2000	personnalisé [3], pas de chevauchement fréquentiel	11 ME avec une surdité sévère à profonde	reconnaissance des mots dissyllabiques et des monosyllabes sans significations	NT, T	- 5/11 sujets préfèrent la compression mais elle n'améliore pas toujours les performances	- la compression n'améliore pas nécessairement la compréhension des sujets; elle peut être une bonne alternative avant l'implantation cochléaire
McDermott & Knight, 2001	AVR ImpaCt	3 ME	reconnaissance de mots monosyllabiques CNC, reconnaissance des consonnes dans un contexte /a/-C-/a/, intelligibilité des phrases dans le bruit (RSB = +10 dB)	NT, T	- pour les tests de reconnaissance des CNC et des consonnes : T = NT; - pour la compréhension dans le bruit T < NT	- la transposition fréquentielle ne fournit pas le bénéfice escompté; - la transposition fréquentielle n'était probablement pas adaptée pour les patients testés et des études complémentaires sont nécessaires pour mieux cibler les indications de la transposition
Simpson et al., 2005	SHARP + Phonak supero 412	17 ME	reconnaissance des phonèmes, des consonnes, des fricatives, des voyelles	NT, T	- pour 8/17 sujets T > NT; - pour 8/17 sujets T = NT; - pour 1/17 sujets T < NT	- globalement il y a une amélioration significative des performances avec la compression mais les résultats ne sont pas uniformément distribués à travers les sujets

Simpson et al., 2006	SHARP + Phonak supero 412	7 ME avec une surdité à pente abrupte (surdité profonde au-delà de 1500 Hz	reconnaissance de mots monosyllabiques CNC, reconnaissance des consonnes dans un contexte vCv, intelligibilité des phrases dans le bruit (recherche du seuil à 50%)	NT, T	<ul style="list-style-type: none"> - pour la reconnaissance des mots et des consonnes T = NT; -3 sujets ont de meilleures performances avec T mais la différence n'est significative que pour un sujet; - 2 sujets ont de moins bonnes performances avec T et la différence est significative pour 2 des 3 sujets; - pour l'intelligibilité dans le bruit T > NT mais cette différence est due au score d'un seul sujet 	- les auteurs supposent qu'il y a une limite sur la quantité d'information qui peut être comprimée dans l'audition encore exploitable
Baskent & Shannon, 2007	vocodeur pour reproduire compression, expansion et décalage fréquentiel	5 NE	reconnaissance des consonnes et des voyelles	NT, T pour différents traitements de compression -expansion et de décalage fréquentiel	<ul style="list-style-type: none"> - compression et expansion < référence; - décalage fréquentiel < référence 	- les performances sont les moins dégradées lorsqu'il y a une compression fréquentielle couplée à un décalage des fréquences vers les graves
Gifford et al., 2007	AVR Nano Xp	6 ME (seuils < 60 dB jusqu'à 500 Hz et > 80 dB à 2000 Hz et au-delà	reconnaissance des mots monosyllabiques CNC, reconnaissance des phrases dans le silence et dans le bruit (RSB = +10 dB)	NT, T	T = NT pour toutes les mesures testées	- la compression fréquentielle ne fournit pas d'amélioration significative chez cette population de sourds partiels

Kulkarni et al., 2007 et 2009	personnalisé [4], pas de chevauchement fréquentiel	6 NE en condition de perte auditive simulée et 11 ME (perte moyenne à profonde)	reconnaissance des mots monosyllabiques type CVC, dans le silence (NE et ME) et dans le bruit (NE)	NT, T avec des ratios de compression de 0,8; 0,6; 0,4	- chez le NE, pour RSB positif NT > T; - pour RSB négatif T > NT (+17% avec T pour les RSB < -6 dB); - chez le ME, T > NT pour la majorité des sujets, pour les diverses conditions. - Pour un ratio de 0,6, l'amélioration avec T est significative chez tous les sujets	- les facteurs de compressions 0,4 et 0,8 fournissent des performances intermédiaires à ce qui est obtenu pour les facteurs 1 et 0,6 ; - il est important de savoir maîtriser les paramètres de la compression, une compression trop importante (0,4) peut nuire à la compréhension par rapport à un ratio optimal (0,6)
Robinson et al., 2007 et 2009	personnalisé [5], avec chevauchement fréquentiel, testé en laboratoire (2007) puis embarqué dans une aide auditive Phonak Aero 311 (2009)	7+5 ME avec des pertes sévères à profondes dans les hautes fréquences, associées à des ZICs	reconnaissance des consonnes en contexte VCV, discrimination de 's' et 'z' en fin de mot, reconnaissance des phrases dans le bruit (2009)	NT, T	- seule la perception des consonnes affriquées est améliorée par la transposition ; - 2007: pour 2/7 sujets T > NT - 2009: pour 2/5 sujets T < NT.	- la perception des fricatives ('s' et 'z') est améliorée par la transposition (seulement pour l'étude de 2007) alors que les performances générales ne le sont pas --> l'amélioration de la détection des fricatives se fait au détriment de la perception des autres catégories de consonnes; - on peut supposer que la transposition fréquentielle n'est adaptée que pour une minorité de patients; - l'algorithme de transposition testé fournit un résultat équivalent à un simple filtrage passe-bas
Xiao et al., 2008	un algorithme M1 à compression avec chevauchement fréquentiel vs. un algorithme M2 sans chevauchement	6 ME avec une perte moyenne à sévère	reconnaissance des consonnes, des mots monosyllabiques, des mots ou des phrases courtes, dans le silence ou dans un bruit blanc à un RSB de +5 dB	NT, M1, M2	M2 > M1 > NT	- en milieu bruyant, les systèmes à compression sans chevauchement fréquentiel sont plus stables que les systèmes avec un chevauchement fréquentiel

Glista et al., 2009	compression non linéaire de fréquence embarqué dans des contours similaires aux Phonak Savia 311 ou 411	24 ME (13 adultes + 11 enfants) avec une surdit� sév�re � profonde dans les hautes fréquences	d�tection des sons de parole (/s/ et /sh/), reconnaissance des consonnes, reconnaissance du pluriel pour d�tecter les /s/ et /z/ en fin de mot, reconnaissance des voyelles	NT, T	- l'ensemble des tests indiquent que T > NT, mais ce n'est pas valable � l'�chelle individuelle ; - le score des voyelles n'est pas affect� par T	- contrairement � l'algorithme test� dans l'�tude de Robinson et al. (2007), l'am�lioration de la reconnaissance des consonnes ne nuit pas en contrepartie � la perception d'autres �l�ments de la parole tels que les voyelles
Bohnert et al., 2010	SoundRecover	14 ME, la plupart avec une perte s�v�re � profonde	reconnaissance des phrases dans le bruit (recherche du seuil � 50%) (11/14 sujets)	NT, T	- pour 7/11, l'intelligibilit� dans le bruit est am�lior�e avec T (de 0,3 dB � 7,35 dB d'am�lioration); - pour les 4 autres sujets T < NT (jusqu'� -7,85 dB)	- tous les patients ne tirent pas un b�n�fice de la compression fr�quentielle

Tableau I. R capitulatif des principales  tudes de ces 50 derni res ann es et des r sultats qui en  mergent.

La littérature sur la compression fréquentielle observe une très large disparité des résultats. Certains patients, pourtant candidats à la compression d'après leur audiogramme, n'arrivent pas à tirer profit de ces systèmes. Certains des échecs observés peuvent être liés à la technologie mais aujourd'hui les appareils auditifs permettent de réduire les distorsions du son par rapport au signal original ; nous constatons encore des échecs malgré cela. Ces patients ne tirent aucun bénéfice de l'amplification conventionnelle ni de la compression fréquentielle. D'autre part ils entendent encore trop bien pour être candidats à l'implant cochléaire. Quelle alternative peut-on leur proposer ?

I.9. La Stimulation Electrique-Acoustique (EAS)

Principe de l'EAS

La Stimulation Electrique-Acoustique (EAS = Electric-Acoustic Stimulation) consiste en une stimulation bimodale du système auditif : une aide auditive amplificatrice va stimuler les basses fréquences acoustiquement tandis que les hautes fréquences seront stimulées électriquement par un implant cochléaire inséré partiellement dans la zone non fonctionnelle de la cochlée, généralement dans le premier tour basal, soit 18 à 20 mm d'insertion du porte-électrodes, ce qui correspond à une zone de fréquences de 700 – 1000 Hz pour l'électrode la plus apicale. Toutefois cette stimulation bimodale n'est possible que si l'audition résiduelle a été préservée post-implantation.

Le challenge pour l'EAS : la préservation de l'audition post-implantation

Dans les années 1990s, les travaux de Shepherd et de son équipe ont mis en évidence, chez l'animal, qu'il était possible de préserver l'audition dans une oreille implantée. Ils ont ainsi démontré que, s'il n'y a pas de trauma d'insertion de l'électrode, d'infection ou d'inflammation sévère, les cellules ciliées apicales au faisceau d'électrodes implanté peuvent non seulement survivre, de plus elles peuvent également fonctionner de façon quasi normale (Shepherd et al., 1983 ; Ni et al., 1992 ; Xu et al., 1997).

En 1997, Hodges et al. ont montré, chez l'humain, qu'il était possible de conserver l'audition résiduelle suite à une implantation cochléaire. Ils ont réalisé pour cela une audiométrie oreilles nues de 40 sujets avant et après leur implantation. Sur les 40 sujets, 21 (52%) ont eu une conservation de leur audition résiduelle et sur ces 21 sujets qui avaient eu leur audition préservée, 76% avaient des seuils audiométriques conservés à 10 dB près par rapport à leurs seuils préimplantation. Sur le groupe des 40 sujets testés, la chute moyenne de l'audition était de 12 dB dans l'oreille implantée et de 4,4 dB dans l'oreille contrôle. Pour autant les 21 sujets qui avaient eu leur audition conservée n'avaient pas de meilleures performances avec leur implant par rapport au groupe de 19 sujets qui avaient perdu l'audition.

Quelques années plus tard, en 1999, Von Ilberg et al. ont introduit le concept de l'EAS en montrant qu'il était possible de préserver et d'exploiter acoustiquement l'audition résiduelle d'un patient implanté cochléaire. Cette expérience chez l'humain a ensuite été renouvelée dans d'autres centres par Kiefer et al. (2002), Skarzynski et al. (2002), Gantz et Turner (2003), Kiefer et al. (2004), James et al. (2006) et Fraysse et al. (2006) qui ont confirmé qu'il est possible de préserver l'audition après une implantation cochléaire. Présentée seule, la stimulation acoustique des très basses fréquences fournit pourtant peu d'intelligibilité (Kong et al., 2005 ; Chang et al., 2006, Zhang et al., 2010). Mais lorsque cette information basse fréquence est couplée à la stimulation de l'IC, les performances des patients pour la compréhension dans le bruit s'améliorent drastiquement, jusqu'à 10 dB en RSB par rapport à l'IC classique (Turner et al., 2004 ; Kong et al., 2005 ; Dorman et al., 2008).

Les centres d'implantation portent un intérêt tout particulier au développement de l'EAS, mais la condition préalable pour en bénéficier est une audition résiduelle suffisante post-implantation. La préservation de l'audition est un objectif majeur, c'est pourquoi nous constatons actuellement un développement important des porte-électrodes atraumatiques et des techniques de chirurgie mini-invasive (soft surgery).

Depuis une dizaine d'années, nous avons vu apparaître des nouveaux prototypes de porte-électrodes qui se sont révélés plus ou moins adaptés et plus ou moins efficaces. En 2003, Gantz et Turner ont essayé l'EAS chez 6 patients, avec des implants ayant des faisceaux d'électrodes plus courts que la normale, le but étant de ne pas endommager les cellules apicales lors de la chirurgie. Ils utilisaient un implant Hybrid CI24 (Cochlear) avec 2 types de porte-électrodes, l'un mesurait 6 mm et l'autre mesurait 10 mm. Les 2 types d'implants

avaient 6 canaux et chacun a été implanté chez 3 patients ayant de l'audition résiduelle dans les graves. Les audiométries réalisées 12 mois après l'implantation ont montré que l'audition avait été préservée chez les 6 sujets. Cependant l'utilisation de porte-électrodes très courts (6 mm) est très controversée. D'une part un faisceau de 6 mm ne va probablement pas assez profond dans la cochlée pour stimuler correctement les fibres nerveuses et les neurones du ganglion spiral encore fonctionnels (Woodson et al., 2010). D'autre part une insertion trop courte risque d'être handicapante pour le patient, dans la mesure où il risque d'y avoir une distorsion importante en terme de place tonotopique (Gantz et Turner, 2003 ; Woodson et al., 2010). Il faut également avoir conscience qu'un porte-électrode très court va impliquer une diminution des espaces inter-électrodes, ce qui va favoriser les interactions (Fu et Shannon, 1999). Gantz et son équipe ont donc poursuivi leurs essais cliniques, essentiellement avec l'implant Hybrid CI24 de longueur 10 mm et ont réussi à préserver l'audition, totalement ou partiellement, chez 46 patients parmi 48 implantés (Gantz et al., 2006).

En parallèle, le développement des chirurgies atraumatiques en Europe se fait avec des insertions un peu plus profondes que celles recommandées par Gantz et son équipe. En effet il faut envisager la possibilité que le patient puisse perdre l'audition post-implantation et des travaux ont montré que, dans ce cas de figure, un implant de 10 mm n'était pas aussi efficace qu'un implant de longueur standard et qu'une réimplantation serait plus appropriée (Fitzgerald et al., 2008). Il est donc recommandé d'insérer le faisceau sur un minimum de 360° (environ 18-22 mm, variable en fonction de l'anatomie de la cochlée), ce qui est un bon compromis pour 1/ préserver l'audition résiduelle et 2/ obtenir de bonnes performances avec l'implant seul au cas où l'audition résiduelle serait perdue post-implantation (Gstoettner et al., 2009). Récemment, avec une telle profondeur d'insertion, l'équipe de Gstoettner (2009) a présenté un taux de 100% de préservation de l'audition. Mais généralement, que ce soit pour les implants Cochlear avec des électrodes pérимодиоlaіres et des insertions de 17 mm (James et al., 2006) ou bien pour les implants Med-El avec des électrodes droites et des insertions de 18-24 mm (Gstoettner et al., 2006), le taux de préservation de l'audition est quasi-identique, à savoir 70% des cas. En parcourant la littérature, certains centres disent obtenir jusqu'à 80-90% de préservation de l'audition (Gstoettner et al., 2004 ; Kiefer et al., 2004 ; Kiefer et al., 2005 ; Skarzynski et al., 2006), mais tous les centres n'ont pas les mêmes critères en terme de préservation de l'audition (certains vont considérer qu'il y a préservation si la différence seuils postop – seuils préop est inférieure à 10 dB, d'autres si la différence est inférieure à

30 dB), ce qui peut biaiser les résultats. Pour maximiser les chances de préserver l'audition, il est nécessaire de mieux maîtriser les techniques chirurgicales.

L'EAS serait donc une alternative pour les patients qui sont en échec avec l'amplification conventionnelle et la compression fréquentielle. Dans une précédente étude, nous avons montré que l'appareillage conventionnel fournissait de meilleurs résultats que l'implant cochléaire. Par analogie, nous pouvons nous demander si la compression fréquentielle fournira de meilleurs résultats que l'EAS. Cette question sera traitée dans la section suivante.

I.10. Restitution de l'audibilité des fréquences supérieures à 1000 Hz dans le cas de surdités partielles

(article publié dans les Cahiers de l'Audition : Seldran F, Thai-Van H, Truy E, Beliaeff M, Berger-Vachon C, Collet L, Gallégo S. 2010 ;4 :24-33)

Restitution de l'audibilité des fréquences supérieures à 1000 Hz dans le cas de surdités partielles

Fabien Seldran^{1,2}, Hung Thai-Van¹, Eric Truy¹, Michel Beliaeff², Christian Berger-Vachon¹, Lionel Collet¹, Stéphane Gallégo^{1,3}

1. Laboratoire Neurosciences Sensorielles, Comportement, Cognition. CNRS UMR 5020. Université Claude Bernard Lyon I.

Pavillon U, Hôpital Ed. Herriot - 3 Place d'ARSONVAL - 69437 Lyon CEDEX 03

2. Vibrant MED-EL Hearing technology - 400 Av Roumanille - BP 309 - 06906 Sophia-Antipolis CEDEX

3. Audition Conseil - 22 Rue Constantine - 69001 Lyon

Résumé

L'implant cochléaire reste un dilemme pour les malentendants atteints de surdités supérieures à 80 dB au-delà de 1000 Hz. Les patients qui ont participé à cette étude présentent ce type de perte ; ils ne tirent pas de bénéfice d'une aide auditive conventionnelle. Malgré tout ils perçoivent encore « trop bien » les sons de basse fréquence et une implantation cochléaire risquerait de détruire leurs restes auditifs.

Aujourd'hui, deux systèmes peuvent être proposés, l'appareil Naïda qui est une aide auditive amplificatrice qui intègre une fonction de compression non-linéaire de fréquence et l'implant à Stimulation Electrique-Acoustique (EAS) de Med-El.

Cependant la littérature reporte des « échecs » inexplicables de chacun de ces systèmes et nous pensons qu'il est important de mieux redéfinir les indications de ces systèmes, afin de choisir l'appareillage qui sera le mieux adapté et de maximiser les chances de réussite du patient. Nous proposons d'utiliser un test d'audiométrie vocale filtrée qui pourra être utilisé par la suite dans le but de réaliser un bilan pronostique qui nous dira « quel système est le mieux adapté à quel profil de patient ».

Le système à compression fréquentielle a pour avantage majeur de supprimer le phénomène de larsen qui est quasi-inévitable avec une amplification linéaire.

L'implant EAS, quant à lui propose d'exploiter les reliquats d'audition dans les basses fréquences et d'utiliser un implant cochléaire pour prendre le relais de la stimulation acoustique.

Nous avons testé le système Naïda sur 5 patients et l'implant EAS sur 4 patients.

Les deux systèmes testés nécessitent un gros travail de rééducation auditive, particulièrement l'implant EAS. Globalement, les résultats avec l'implant EAS sont plutôt positifs.

1

Introduction

La présente étude porte sur l'appareillage des surdités de perception type « pente de ski », c'est-à-dire une perte auditive importante prédominant sur les fréquences aiguës. Les patients concernés ont une bonne perception dans les fréquences graves mais une perte supérieure ou égale à 80 dB HL sur les fréquences supérieures à 1 kHz.

Les études montrent que l'amplification conventionnelle des hautes fréquences a un rôle limité pour les personnes

qui présentent ce profil de surdité (Ching et al, 1998). Turner (2006) montre que l'amplification conventionnelle des hautes fréquences ne fournit une amélioration que si la perte auditive est inférieure à 60-80 dB HL. Elle n'améliore pas l'intelligibilité et parfois elle la dégrade. Généralement, ces patients ont des scores d'intelligibilité inférieurs à 60% avec une amplification conventionnelle. De plus, Vickers et al (2001) ont montré que si le patient présente des Zones Inertes Cochléaires (ZIC), l'amplification ne peut fournir un bénéfice que si elle est limitée en fréquence jusqu'à 1,5 à 2 fois la fréquence de coupure (Fzic). Pour ces personnes, les hautes fréquences ne sont pas exploitables acoustiquement. En général, les aides auditives fournissent donc des résultats non satisfaisants, en particulier pour la perception de la musique et pour la compréhension dans le bruit. Les systèmes d'implants cochléaires ne sont pas non plus adaptés pour ces patients car ils ne pourront pas bénéficier de leurs reliquats auditifs dans les basses fréquences.

Actuellement, deux alternatives peuvent être proposées pour restituer les hautes fréquences lorsque l'amplification conventionnelle atteint ses limites et que le patient n'en tire qu'un maigre bénéfice : les systèmes amplificateurs à compression ou transposition fréquentielle et les implants à stimulation électrique-acoustique (EAS = Electric-Acoustic Stimulation). Nous proposons dans cette étude d'évaluer et de comparer chez ce type de patients, le système à compression fréquentielle Naïda de Phonak et l'implant électrique-acoustique (Processeur DUET + électrode FlexEAS) de Med-El.

1.1. Etat des connaissances sur la Transposition / Compression fréquentielle

Pour l'appareillage de ces pertes, une alternative avant l'implantation cochléaire serait l'utilisation d'aides auditives (amplificatrices) à transposition ou à compression fréquentielle, ces systèmes ont pour principe de comprimer le spectre de la parole dans la gamme de fréquences encore audibles par le malentendant. Ceci implique généralement une diminution, voire une annulation de larsen (Joson et al, 1993), ce qui est un avantage non-négligeable par rapport aux systèmes d'amplification standard qui y sont très sensibles lorsque le gain dans les fréquences aiguës devient trop important, ce qui est souvent le cas pour ce type de surdité (pente de ski). La théorie de ces systèmes à compression fréquentielle suppose que l'enveloppe temporelle soit préservée et le malentendant perçoit les sons de façon plus grave et un peu plus naturelle. McDermott et

al (1998) et Thai-Van et al (2002 et 2003) ont montré qu'il existait, chez les patients avec une perte importante dans les hautes fréquences, une meilleure discrimination fréquentielle et une sur-représentation des neurones, au niveau cortical, de la dernière zone de fréquences normalement codées. Sur le concept, le malentendant devrait donc être capable d'intégrer l'information transposée par les systèmes à transposition ou à compression fréquentielle, cependant il existe un problème de résolution fréquentielle au niveau de la cochlée (mécanique cochléaire liée aux filtres auditifs de cochlée pathologique).

Concernant les paramètres des systèmes de compression, la littérature propose certaines règles à respecter. Turner et Hurtig (1999) ont montré que le ratio de compression ne devait pas dépasser une valeur de 1,4 et que l'information de parole dans les hautes fréquences doit être décalée dans une zone où la sensibilité est meilleure que 60 dB HL. Un autre paramètre à considérer est la période d'accoutumance au système. D'après Reed et al. (1983), le pattern du son généré par la diminution de fréquence est différent de la parole naturelle non déformée, il est raisonnable de supposer qu'un entraînement, ou une rééducation auditive assez longue est indispensable pour apprendre à « encoder » (à comprendre) les signaux de parole diminués en fréquence.

Baskent et Shannon (2007) ont évalué les effets combinés du décalage fréquentiel et de la compression-expansion sur la reconnaissance de la parole, en utilisant un vocoder à bandes de bruit. Cette simulation est testée sur 5 sujets normo-entendants. Les résultats montrent que lorsqu'il y a des distorsions spectrales de compression-expansion ou de décalage fréquentiel appliquées séparément, les performances sont moins bonnes que dans la condition de référence. Mais lorsque les deux dégradations spectrales sont appliquées simultanément, un effet compensatoire est observé et la diminution des performances est moins importante. Les résultats montrent que les performances sont les moins dégradées lorsqu'il y a une compression fréquentielle couplée à un décalage des fréquences vers les graves (de 3 à 5 mm vers l'apex sur le modèle simulé).

Une étude de Xiao et al (2008) a montré qu'en milieu bruyant, les systèmes à compression sans chevauchement fréquentiel sont plus stables que les systèmes avec un chevauchement fréquentiel. Les systèmes à chevauchement fréquentiel sont des systèmes pour lesquels une bande de fréquence contient une information amplifiée linéairement et une information transposée depuis une bande de fréquence supérieure. Mais globalement, la littérature présente des résultats très mitigés sur le bénéfice des systèmes à compression fréquentielle (Simpson, 2009).

L'ensemble des études passées en revue montre une amélioration du gain prothétique tonal, mais les résultats montrent surtout que le fait de donner accès à des sons de haute-fréquence, non audibles initialement par les sujets, n'engendre pas forcément de meilleures performances en compréhension.

1.2. Etat des connaissances sur la Stimulation Electrique-acoustique

Si un patient a des reliquats d'audition dans les basses fréquences et qu'il ne tire aucun bénéfice de l'appareillage acoustique, ni par les systèmes conventionnels, ni par les

systèmes à compression fréquentielle, les études récentes (Gifford et al, 2007) proposent alors la Stimulation Electrique-Acoustique (EAS) comme une nouvelle alternative pour ces patients. En 1999, Von Ilberg et al ont introduit le concept de l'EAS en montrant qu'il était possible de préserver et d'exploiter acoustiquement l'audition résiduelle d'un patient implanté cochléaire. Cette expérience chez l'humain a ensuite été reproduite notamment par Skarzynski et al (2002) et Gantz et Turner (2003) qui ont confirmé qu'il est possible de préserver l'audition après une implantation cochléaire. Von Ilberg et al (1999) ont en particulier montré que la combinaison d'une stimulation acoustique et électrique n'engendrait pas d'interférence perturbante et que le système auditif central était capable d'intégrer cette stimulation bimodale. Kiefer et al (2006) ont montré que la présence de l'électrode ne changeait pas le comportement de la membrane basilaire (à l'apex) et que les basses fréquences n'étaient pas affectées par la fixation de la membrane basilaire sur le tour basal et milieu de la cochlée.

L'EAS consiste en une stimulation bimodale du système auditif ; une aide auditive amplificatrice va stimuler les basses fréquences acoustiquement tandis que les hautes fréquences seront stimulées électriquement par un implant cochléaire inséré partiellement dans la zone non fonctionnelle de la cochlée, usuellement dans le premier tour basal, soit 18 à 20 mm d'insertion du porte-électrodes, ce qui correspond à une zone de fréquences de 700 - 1000 Hz pour l'électrode la plus apicale.

Depuis quelques années, des nouvelles techniques de chirurgie mini-invasive (Briggs et al, 2006 ; Lenarz et al, 2006), couplées à l'utilisation de nouveaux faisceaux d'électrodes permettent de minimiser le traumatisme induit par l'acte chirurgical et ainsi de ne pas détruire l'audition résiduelle, comme c'est le cas pour l'implantation cochléaire conventionnelle (James et al, 2006 ; Gantz et Turner, 2003). Pour l'EAS il est préconisé de faire une insertion partielle du porte-électrode dans la cochlée, cela suppose qu'on va préserver une partie de l'audition dans les basses fréquences. Gantz et Turner (2003) qui ont implanté 3 patients avec une profondeur d'insertion de 6 mm et 3 patients avec une profondeur d'insertion de 10 mm, constatent que l'audition est partiellement préservée chez les 6 sujets (perte de 10-15 dB) et ils constatent surtout que les scores de perception de la parole préopératoires sont inchangés après la chirurgie. Cependant une insertion partielle, bien qu'elle permette de préserver l'audition dans les basses fréquences, implique néanmoins la destruction d'une autre partie de l'audition dans les hautes fréquences. Adunka et Kiefer (2006) montrent que la sévérité du trauma intracochléaire augmente pour les insertions profondes (dès qu'on enfonce le porte-électrode au-delà du point de 1ère résistance).

Les résultats de la littérature montrent un effet synergétique de la combinaison des deux stimulations et ils précisent qu'il faut compter plusieurs mois avant que le gain en intelligibilité soit apparent (Gantz et Turner, 2003). Le fait de pouvoir utiliser l'audition résiduelle dans les basses fréquences permet aux patients de mieux apprécier la musique, d'améliorer la compréhension dans le bruit (Gantz et Turner, 2003 ; Kiefer et al, 2005 ; Dorman et al, 2005 ; Turner et al, 2004) et surtout d'offrir une sensation d'écoute plus naturelle comparée à un implant cochléaire classique.



Figure 1. Photographie de l'aide auditive Naïda

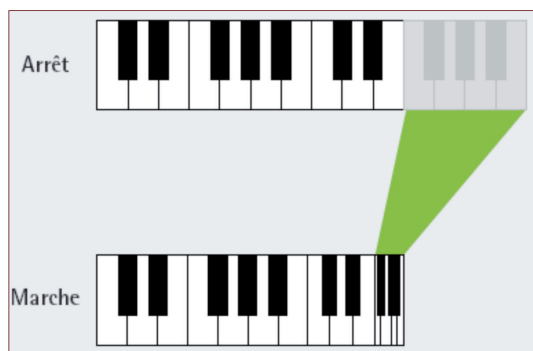


Figure 2. Schéma de principe de la compression fréquentielle du Naïda (Source Phonak).

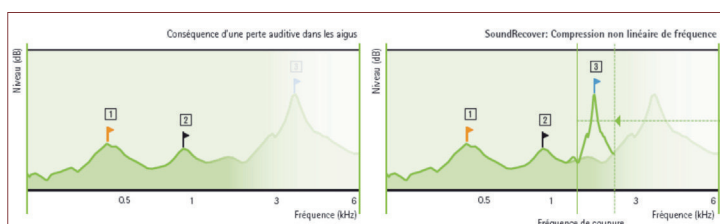


Figure 3. Représentation spectrale du traitement du signal réalisé par le Naïda (Source Phonak). A gauche, le spectre d'un signal perçu par un malentendant sans aide auditive. A droite, le spectre du même signal perçu par un malentendant appareillé avec un système à compression fréquentielle. Le spectre dans les hautes fréquences a une bande plus étroite, décalée vers les fréquences graves.

Les restes dans les basses fréquences fournissent en effet un meilleur accès aux informations de prosodie et de voisement (James et al, 2006) ainsi qu'à la fréquence fondamentale et au premier formant des sons de parole (Kiefer et al, 2005).

Gilbert et Micheyl (2005) ont étudié les fonctions d'importance de différentes bandes de fréquences et ont montré que la gamme de fréquences [100-750] Hz était la moins affectée par le bruit (competitive babble noise) et relativement aux autres bandes, elle devient la bande la plus importante en milieu bruyant par rapport à un milieu calme. Le fait d'utiliser l'audition résiduelle pour l'EAS semble être un atout majeur pour améliorer la compréhension dans les situations bruyantes.

Toutefois l'EAS ne peut être effective que si l'audition résiduelle a été préservée et actuellement, le risque de perdre l'audition suite à l'implantation n'est pas nul puisque les études montrent que l'audition est partiellement préservée dans 70 à 90% des cas (James et al, 2006 ; Kiefer et al, 2005 ; Gstoettner et al, 2006 ; Skarzynski et al, 2006 ; Gstoettner et al, 2004 ; Kiefer et al, 2004). Actuellement, l'objectif majeur de la recherche sur l'EAS est de déterminer les paramètres qui permettent de minimiser les risques de destruction des reliquats, mais également de déterminer si un patient va être un bon candidat pour l'EAS ou si au contraire, l'EAS risque de ne lui fournir qu'un bénéfice limité. De plus en plus d'études se focalisent sur la technique chirurgicale, particulièrement sur le choix d'une insertion de l'électrode par la fenêtre ronde versus cochléostomie (Briggs et al, 2006 ; Kiefer et al, 2004) et sur la détermination de la profondeur d'insertion de l'électrode à partir des images de la cochlée avant l'implantation

(Adunka et al, 2005) ou à partir de l'équation de Greenwood (Kiefer et al, 2004).

La littérature sur l'EAS montre que certains candidats ont de bonnes performances avec l'EAS. Cela signifie que la méthode fonctionne et qu'elle est validée. Mais on constate également des échecs et tous ne sont pas expliqués. Il reste donc encore des inconnues sur l'EAS et c'est ce que nous essayons de mieux appréhender.

Une partie des échecs de l'EAS sont dus au fait que l'audition résiduelle a été détruite, tandis que pour d'autres cas l'EAS ne fournit aucun bénéfice, bien que l'audition résiduelle soit partiellement préservée après la chirurgie (Gstoettner et al, 2008 ; Kiefer et al, 2005 ; Gantz et al, 2005).

Ces différents échecs montrent donc que malgré la préservation de l'audition, l'EAS n'est pas toujours bénéfique, ceci infirme les propos de Von Ilberg et montre que tous les patients n'ont pas un système auditif central capable d'intégrer la stimulation bimodale.

1.3. Bilan sur l'état des connaissances

Que ce soit pour les systèmes amplificateurs à transposition fréquentielle ou à compression fréquentielle, ou que ce soit pour les implants EAS, nous savons que certaines techniques fonctionnent chez certains patients mais pas chez d'autres. Des échecs sont constatés et inexpliqués. Nous supposons alors qu'il existe plusieurs profils de patients. Etant donné qu'il y a des échecs de la compression fréquentielle et de l'EAS chez certaines personnes, cela signifie que le bilan n'est pas adapté avec ces systèmes.

Ces deux types de systèmes ont pour principe de réhabiliter les hautes fréquences acoustiques. Le bilan doit donc évaluer l'intelligibilité de ces hautes fréquences. Pour prédire le résultat obtenu avec ces systèmes, nous souhaitons investiguer la contribution fréquentielle dans l'intelligibilité du malentendant. Nous pourrions savoir si l'ajout des hautes fréquences apporte de l'intelligibilité. Nous pourrions également définir différents profils de patients afin de déterminer, de façon plus précise quel appareillage est le mieux adapté à quel profil de patient.

Dans ce but, nous avons utilisé une batterie de tests audiométriques simples (suggérés par Turner, 2006) que nous réalisons avec les sujets inclus dans notre protocole de tests.

1.4. Les systèmes étudiés

1.4.1. Le système Naïda

Le système Naïda de Phonak (figure 1) est un contour numérique surpuissant qui intègre une fonction de compression non-linéaire de fréquence : le Soundrecover. La compression fréquentielle est destinée à des malentendants qui ne perçoivent plus les sons aigus. La compression fréquentielle repose sur deux principes. Le premier principe consiste à comprimer des bandes de fréquences larges dans les hautes fréquences (inaudibles par le patient) dans une bande de fréquence moins large. Le second principe consiste à ramener cette zone de fréquences comprimées depuis une zone inaudible vers une zone où le patient perçoit encore les sons. C'est-à-dire que les composantes fréquentielles des bandes comprimées seront toutes diminuées en

fréquence, par un même ratio. L'enveloppe temporelle du signal n'étant pas dégradée, le signal de sortie sera simplement perçu plus grave que le signal d'entrée.

Pour imaginer ce concept, supposons que nous jouons un morceau de musique sur un piano (**figure 2**). Si le morceau est joué dans les aigus, il ne pourra pas être perçu par le malentendant. Si ensuite le même morceau est joué une octave en dessous sur le piano (donc dans les sons plus graves), il pourra être perçu par le malentendant car les sons seront audibles pour lui. L'aspect temporel du signal sera inchangé, la mélodie restera la même, seul l'aspect fréquentiel changera puisque le « timbre » perçu sera plus grave que l'original.

Pour résumer, les bandes de basse fréquence ne sont pas comprimées jusqu'à une certaine fréquence (F_c). Au-delà de cette fréquence F_c , les bandes perçues sont des bandes comprimées et transposées vers des fréquences plus graves (**figure 3**). Toutes ces bandes sont adjacentes et ne se chevauchent pas en fréquence. La compression du spectre dans les hautes fréquences engendre ainsi une distorsion moins importante que la compression de la totalité du spectre. Ceci permet de préserver les indices importants de la parole dans les basses fréquences.

Le soundrecover propose de choisir le ratio de compression ($[\text{largeur de la bande initiale}] / [\text{largeur de la bande comprimée}]$) ou la fréquence de démarrage de la compression. Ratio de compression et fréquence de démarrage évoluent en parallèle puisqu'ils dépendent mutuellement l'un de l'autre dans le soundrecover.

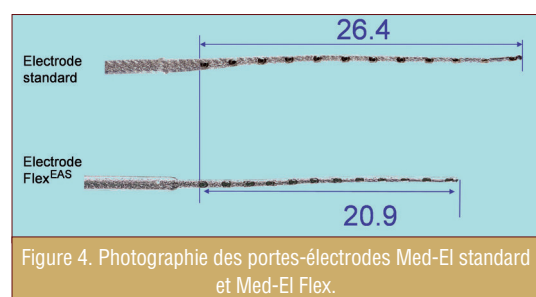


Figure 4. Photographie des portes-électrodes Med-El standard et Med-El Flex.

1.4.2. Le système DUET de MED-EL

Le système EAS de Med-El est le seul homologué et disponible actuellement sur le marché, il intègre les technologies de l'implant cochléaire et de l'aide auditive numérique. Ce système EAS est composé d'un porte-électrodes FlexEAS et d'un processeur DUET EAS.

Le porte-électrodes FlexEAS a été spécialement conçu, de façon à protéger au mieux les structures fragiles de la cochlée contre les effets traumatiques de l'insertion du porte-électrodes, car la préservation de l'audition résiduelle dans les basses fréquences est un des facteurs clés de la réussite de l'EAS. Par conséquent, le porte-électrodes doit avoir un design qui permet de minimiser ce traumatisme. En l'occurrence, le porte-électrodes Flex de Med-El mesure 24 mm, il possède 12 canaux sur 20,9 mm avec un espace inter-électrode de 1,9 mm ($11 \text{ espaces} \times 1,9 \text{ mm} = 20,9 \text{ mm}$) et le diamètre de l'électrode à l'extrémité basale est de 0,8 mm. Enfin, le porte-électrodes est relié à un boîtier d'implant SONATA. A titre d'exemple, la figure 4 montre une

photographie du porte-électrodes Flex comparé à un porte-électrodes standard de chez Med-El.



Figure 5. Le processeur externe DUET de Med-El qui intègre une aide auditive et un processeur EAS pour la partie implant cochléaire.

Le processeur DUET EAS (**figure 5**) se présente sous la forme d'un contour d'oreille. Il combine les fonctions d'aide auditive grâce à un module acoustique qui stimule acoustiquement les basses fréquences (125 – 1500 Hz) et une fonction implant cochléaire via un module qui stimule de façon électrique les fréquences aiguës.

Le microphone du DUET va capter les sons de la même façon qu'un contour classique. Le son est ensuite numérisé puis analysé par le processeur vocal qui transmet le signal à deux circuits indépendants en parallèle (**figure 6**).

Le module acoustique est une aide auditive à trimmer dédiée à la stimulation des sons graves (125 - 1500 Hz). Le signal amplifié est transmis au conduit auditif via un embout similaire à celui d'une aide auditive conventionnelle.

Le module implant cochléaire est dédié au traitement des sons medium aigus, le signal va être divisé à travers 10 filtres passe-bande (en considérant que seules 10 électrodes sur 12 sont insérées dans la cochlée), chacun de ces filtres codant le signal d'un canal de l'implant. Comme dans l'implant cochléaire classique, pour chaque canal le signal subit une extraction d'enveloppe par une transformée de Hilbert et l'électrode correspondante stimule le nerf auditif par une impulsion électrique (pulse biphasique équilibré en

charge) proportionnelle à l'enveloppe du signal. L'implant Med-El utilise une stratégie de codage CIS (Continuous Interleaved Sampling = Stimulation Intercalée en Continu), c'est-à-dire que toutes les électrodes sont stimulées, de façon séquentielle et à une cadence rapide.

2

Matériel et méthode

2.1. Population d'étude

20 sujets normo-entendants et 9 patients malentendants ont été inclus dans l'étude globale. Les 20 sujets normo-entendants étaient âgés de 18 à 30 ans, leur seuil d'audition était inférieur à 20 dB HL sur les fréquences 250 - 8000 Hz.

Les 9 malentendants qui ont participé avaient de 30 à 78 ans, avec une moyenne de 54,9 ans. Les 9 malentendants avaient une perte supérieure à 70 dB HL au-delà de 1000 Hz, le seuil moyen dans la gamme 250-710 Hz est de 51 dB HL pour les patients Naïda et 68 dB HL pour les patients EAS ; le seuil moyen dans la gamme 1000-4000 Hz est de 95 dB HL pour les patients Naïda et 111 dB HL pour les patients EAS. Quatre ont été implantés EAS (3 oreilles droites + 1 oreille gauche) et cinq ont été appareillés avec le système Naïda (8 oreilles = 4 oreilles droites + 4 oreilles gauches). Le patient Naïda n°3 est implanté cochléaire, pour ce patient le Naïda est une aide à l'implant. Un des patients implantés EAS a eu une surdité brusque, tous les autres patients ont eu une surdité évolutive.

2.2. Matériel phonétique et matériel de test

Le matériel phonétique utilisé est une série de 40 listes de mots dissyllabiques de Fournier numérisés (40 listes de 10 mots), prononcés par une voix masculine. Des études de Cutler et al (1983 et 1989) ont montré que les francophones utilisent une stratégie de compréhension basée sur la segmentation des mots en syllabes contrairement aux anglophones qui segmenteraient plutôt en phonèmes. Par conséquent nous avons choisi d'utiliser des mots dissyllabiques pour notre test.

Nous utilisons des mots dissyllabiques, c'est pourquoi l'unité de comptage considérée est le nombre de syllabes correctement répétées. Usuellement, l'unité de comptage du test de Fournier est le nombre de mots correctement répétés mais nous avons choisi ici de compter en syllabes afin d'avoir une meilleure précision. Nous avons également fait ce choix, car dans les situations où la parole est très dégradée, les scores avoisinent le 0% de reconnaissance des mots, alors que le nombre de syllabes répétées n'est pas négligeable. Chaque liste contient 10 mots dissyllabiques, donc chaque syllabe juste vaut 5% de reconnaissance.

Les différentes situations testées étaient des signaux filtrés en passe-bas à des fréquences de coupures de 250 Hz; 353 Hz; 500 Hz ; 707 Hz ; 1000 Hz ; 1414 Hz ; 2000 Hz, ainsi qu'une situation où nous diffusons le signal naturel, non filtré.

Tous les filtrages de nos signaux ont été réalisés avec le logiciel Cool Edit Pro (Adobe Audition) grâce à la fonction « filtre FFT ». Les paramètres des filtres que nous avons

choisis étaient des coefficients de 100% dans la bande passante et 0% dans la bande atténuée. 0% correspondant à une atténuation de plus de 70 dB. Toutes les pistes audio ont été gravées sur CD. Nous avons choisi de créer 1 CD pour chaque fréquence de coupure. Nous diffusons ensuite le CD sur le lecteur qui est relié à l'audiomètre, et toutes les stimulations sont diffusées au casque.

2.3. Déroulement et description des tests

2.3.1. Valeurs normatives

La première partie de l'étude a consisté à étudier chez le normo-entendant, la contribution des basses fréquences dans le signal de parole. Des signaux de parole ont été traités par des filtres passe-bas à différentes fréquences de coupure, ceci afin d'établir, chez le normo-entendant, un modèle « pourcentage de reconnaissance des syllabes en fonction de la fréquence de coupure du signal de parole ». Cette étape nous a permis d'obtenir des valeurs normatives pour notre modèle, par rapport à la langue et aux listes de mots que nous avons utilisés.

Cette pré-étude a été réalisée sur notre population de 20 sujets normo-entendants, sur l'oreille droite uniquement, sans entraînement au test. Une étude de Gilbert et al (2003) sur les fonctions d'importance fréquentielle a montré que, chez le normo-entendant, il n'apparaissait pas de différence de perception entre oreille droite et oreille gauche.

Pour chaque fréquence de coupure (chaque CD), les intensités de stimulations sont 10 ; 20 ; 30 ; 40 ; 50 ; 60 ; 70 ; 80 dB SPL du signal original avant filtrage, les pistes filtrées sont donc moins énergétiques que ce qui est affiché sur l'appareil. Toutes les pistes sont diffusées dans un ordre aléatoire, mais l'ordre de passation est le même, à savoir d'abord signal non filtré, puis signal de plus en plus filtré passe-bas, pour des intensités de plus en plus fortes.

2.3.2. Evaluation chez les patients

La seconde partie de l'étude a consisté à réaliser les tests chez les malentendants. Avant l'appareillage, chaque malentendant a été soumis à une audiométrie tonale, par ½ octave entre 250 et 6000 Hz (entre 250 et 4000 Hz pour les patients implantés EAS). Nous avons recherché son seuil de perception et son seuil subjectif d'inconfort, afin d'évaluer son champ auditif résiduel. Ensuite le patient est soumis à une audiométrie vocale avec les listes de Fournier à différentes intensités, à 70 ; 80 ; 90 ; 100 dB SPL, en établissant un score sur 2 listes = 20 mots, soient 40 syllabes. Enfin, nous nous plaçons à l'intensité jugée la plus confortable pour le patient, pour le signal d'origine (généralement entre 90 et 100 dB SPL) et nous restons au même niveau pour reproduire l'audiométrie avec les listes de Fournier filtrées, avec les mêmes stimulations utilisées dans la première partie de l'étude, chez les normo-entendants. Nous conservons ainsi la même amplification par l'audiomètre en changeant seulement la fréquence de coupure. Afin d'avoir des scores plus précis, nous diffusons là encore 2 listes pour chaque situation filtrée, ce qui nous laisse une précision de comptage de 2,5 %. Toutes les listes ont été diffusées aléatoirement et l'ordre de passation du test reste le même quel que soit le patient, à savoir d'abord signal non filtré, puis signal

de plus en plus filtré passe-bas. L'oreille droite et oreille gauche sont testées séparément et par ordre de passation aléatoire. L'évaluation de l'apport fourni par l'appareillage a consisté en une mesure du gain prothétique tonal et du gain prothétique vocal, à J3 semaines (3 semaines après la pose de l'appareil) pour les patients appareillés avec le Naïda et à J3 mois (3 mois après l'activation de l'appareil) pour les patients implantés EAS. Les implantés EAS ont eu besoin de plus de temps pour intégrer la stimulation bimodale (électrique et acoustique). Nous avons considéré que ces délais étaient convenables pour que les patients soient un minimum habitués à la sensation sonore fournie par leur nouvel appareil.

Nous avons ensuite recherché s'il existe une relation entre les performances à l'audiométrie vocale filtrée et le gain prothétique vocal pour tenter de répondre à la problématique « quel type d'appareillage est le mieux adapté à quel profil de patient ? ».

Si notre test d'audiométrie vocale filtrée montre qu'il peut prédire le bénéfice de la compression fréquentielle et de l'implant EAS, nous pourrions l'utiliser dans un but pronostique dans le bilan, afin de mieux redéfinir les indications d'appareillage avec ces deux systèmes.

3

Résultats

3.1. Intelligibilité en fonction de la bande-passante, chez le normo-entendant

Les résultats figure 7a obtenus chez 20 normo-entendants représentent l'évolution du pourcentage de reconnaissance des syllabes en fonction de l'intensité de diffusion, pour les différentes fréquences de coupures passe-bas (Fc). Pour chaque Fc, nous constatons le même genre de courbe de type sigmoïde ; le score augmente progressivement jusqu'à un niveau d'intensité pour lequel il sature (généralement 40 dB SPL). Plus la bande passante augmente vers les fréquences aiguës, plus le score de saturation se rapproche de 100% de compréhension.

A partir de cette série de courbes, nous avons calculé le pourcentage de reconnaissance optimale des syllabes en fonction de la fréquence de coupure du filtre passe-bas (figure 7b) en considérant la moyenne des scores aux intensités 40 et 50 dB SPL. Les résultats obtenus et décrits figure 7b, montrent une évolution quasi-linéaire de l'intelligibilité en fonction du logarithme de la fréquence de coupure du signal de 353 à 2000 Hz. Une analyse par régression linéaire, nous confirme statistiquement cette évolution rectiligne ($N = 6$ de 353 à 2000 Hz ; $R^2 = 0,99$; $p < 0,0001$) ; l'intelligibilité plateau augmente de 35% par octave. D'après notre modèle par extrapolation, le 0% est obtenu pour une fréquence de passe-bas de 290 Hz, le 100% pour 2080 Hz.

3.2. Intelligibilité chez le malentendant

Nous avons évalué les performances après appareillage chez 9 patients, cinq ont un système Naïda et quatre ont un implant EAS.

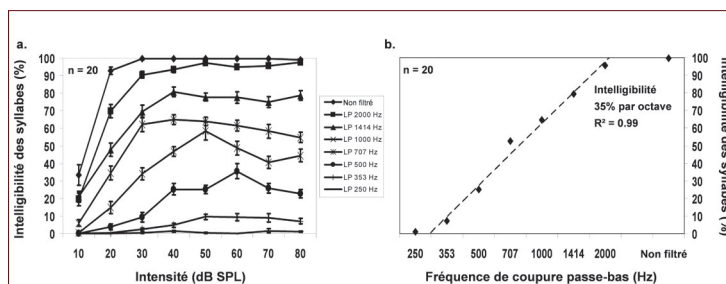


Figure 7a. Pourcentage de reconnaissance des syllabes en fonction de l'intensité de diffusion pour différentes fréquences de coupures du passe-bas. Valeurs moyennes et erreurs standard obtenues chez 20 sujets normo-entendants. Les différentes courbes correspondent aux différentes fréquences de coupure testées.

Figure 7b. Modèle « intelligibilité des syllabes en fonction de la fréquence de coupure ». D'après la moyenne des résultats ($N = 20$ sujets normo-entendants) recueillis aux intensités 40 et 50 dB SPL. La pente est de 35% par octave. La courbe bleue en traits pleins représente notre modèle, la courbe noire en traits pointillés représente la droite de régression linéaire établie à partir des valeurs de notre modèle entre les fréquences de coupures 353 et 2000 Hz.

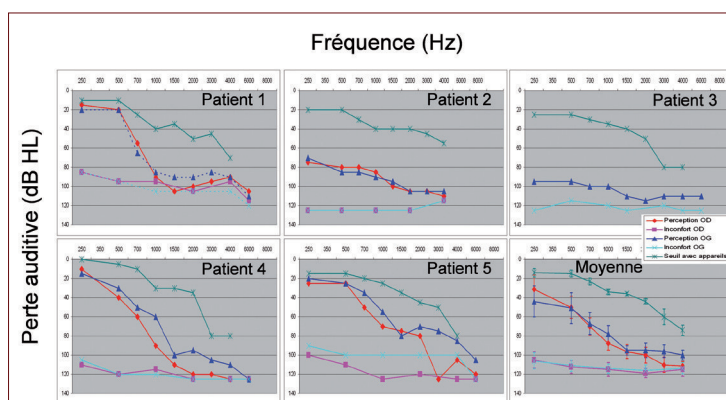


Figure 8. Audiogrammes tonaux des patients appareillés Naïda – seuils de perception et d'inconfort au casque et seuil de perception oreilles appareillées, en champ libre. Une courbe en pointillés signifie que l'oreille testée n'a pas été appareillée. Le graphe en bas à droite représente la moyenne des audiogrammes des 5 patients testés, la barre d'erreur correspond à l'erreur standard.

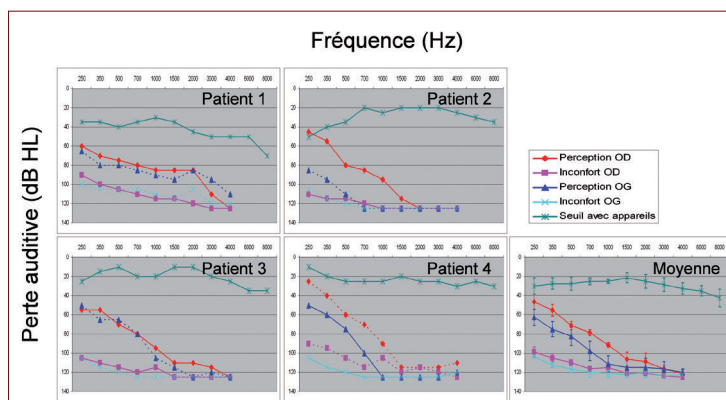


Fig. 9. Audiogrammes tonaux des patients implantés EAS – seuils de perception et d'inconfort au casque et seuil de perception oreilles appareillées, en champ libre. Une courbe en pointillés signifie que l'oreille testée n'a pas été appareillée. Le graphe en bas à droite représente la moyenne des audiogrammes des 4 patients testés, la barre d'erreur correspond à l'erreur standard.

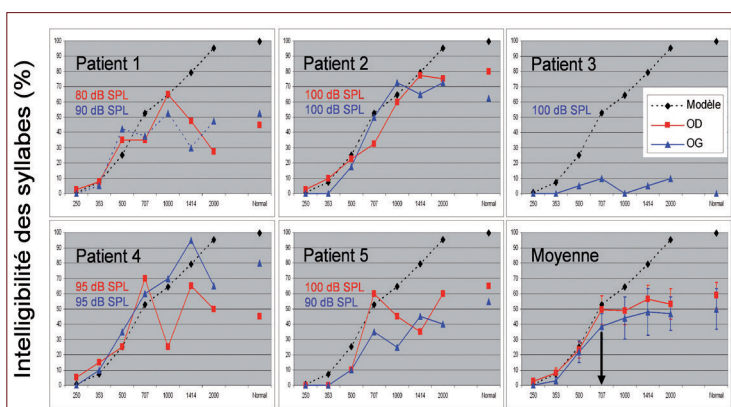


Figure 10. Intelligibilité des syllabes en fonction de la fréquence de coupure chez les patients appareillés en Naïda, avec les listes de Fournier non filtrées (point à l'extrême droite) et filtrées en passe-bas. La courbe noire en pointillés représente le modèle que nous avons établi chez le normo-entendant, dans notre pré-étude. Les chiffres en rouge et bleu indiquent, pour chaque patient, l'intensité de diffusion dans l'oreille droite et dans l'oreille gauche. Une courbe en pointillés signifie que l'oreille testée n'a pas été appareillée. Le graphe en bas à droite représente les résultats moyens des 5 patients testés, la barre d'erreur correspond à l'erreur standard. La flèche noire indique la fréquence au-delà de laquelle nous observons la saturation d'intelligibilité.

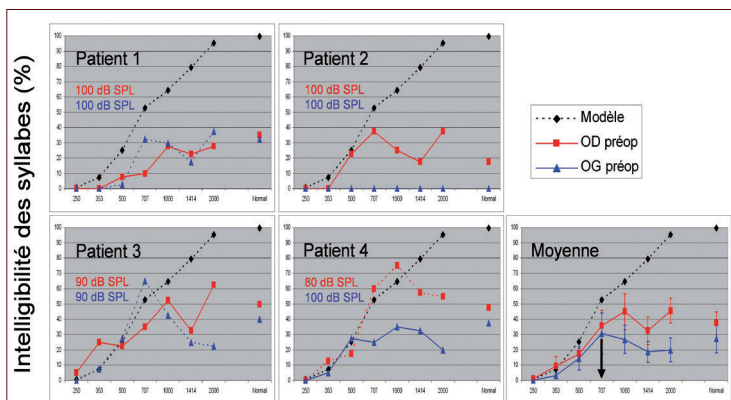


Fig. 11. Intelligibilité des syllabes en fonction de la fréquence de coupure chez les patients implantés EAS, avec les listes de Fournier non filtrées (point à l'extrême droite) et filtrées en passe-bas. La courbe noire en pointillés représente le modèle que nous avons établi chez le normo-entendant, dans notre pré-étude. Les chiffres en rouge et bleu indiquent, pour chaque patient, l'intensité de diffusion dans l'oreille droite et dans l'oreille gauche. Une courbe en pointillés signifie que l'oreille testée n'a pas été appareillée. Le graphe en bas à droite représente les résultats moyens des 4 patients testés, la barre d'erreur correspond à l'erreur standard. La flèche noire indique la fréquence au-delà de laquelle nous observons la saturation d'intelligibilité.

Les graphes, **figures 8 à 13**, représentent les performances individuelles et moyennes :

- à l'audiométrie tonale, seuil de perception (oreille droite = rouge, oreille gauche = bleu), seuil d'inconfort (oreille droite = rose, oreille gauche = bleu ciel) et gain prothétique (vert), pour les patients Naïda et pour les implantés EAS ;
- à l'audiométrie vocale au casque à l'aide de nos listes filtrées (oreille droite = rouge, oreille gauche = bleu, notre modèle = noir), pour les patients Naïda et pour les implantés EAS ;
- à l'audiométrie vocale en champ libre, oreilles nues (orange) vs oreilles appareillées (vert), pour les patients Naïda et pour les implantés EAS.

Pour tous ces graphiques, une courbe en pointillés signifie que l'oreille testée n'a pas été appareillée. Tous les graphiques dans l'angle en bas à droite, représentent la moyenne des résultats obtenus, la barre d'erreur représente l'erreur standard. Pour les patients implantés EAS, les audiométries effectuées au casque et oreilles nues ont été mesurées avant l'implantation. On rappelle que le patient 3 appareillé avec un Naïda porte un implant cochléaire de l'autre côté ; pour ce patient, le Naïda représente une aide à l'implant cochléaire.

* Les performances obtenues à l'audiométrie tonale (figures 8 et 9)

Chez les patients EAS, la perception des différents sons semble équilibrée et les seuils de perception moyens sont situés entre 20 et 40 dB HL et ce jusqu'à des fréquences de 8000 Hz, alors que chez les patients appareillés en Naïda, malgré la puissance de l'appareil et la compression fréquentielle, les patients perçoivent toujours difficilement les sons de haute fréquence.

* Les performances obtenues à l'audiométrie vocale filtrée

Les **figures 10 et 11** représentent les courbes « intelligibilité des syllabes en fonction de la fréquence de coupure ». Nous constatons d'après les différents graphes que, à l'exception du patient Naïda 4, oreille gauche, tous les malentendants que nous avons appareillés ont une courbe qui sature en intelligibilité lorsque la fréquence de coupure augmente, les informations de haute fréquence n'améliorent pas l'intelligibilité de ces patients. La saturation d'intelligibilité (indiquée par la flèche noire dans le graphe en bas à droite des figures 10 et 11) se produit à la fréquence de coupure de 707 Hz.

* Les performances obtenues à l'audiométrie vocale en champ libre

Les performances de compréhension semblent meilleures avec le système Naïda qu'avec le système EAS mais ceci mérite d'être discuté. Si nous regardons les résultats des figures 8 et 9, les patients appareillés en EAS perçoivent mieux les sons de haute fréquence que les patients appareillés en Naïda. D'après les audiométries tonales avec appareils, on pourrait supposer que les patients EAS ont une meilleure compréhension que les patients qui utilisent la compression de fréquences et les résultats montrent que ce n'est pas le cas (**cf figures 12 et 13**).

4

Discussion

4.1. Intelligibilité en fonction de la bande-passante, chez le normo-entendant

La courbe d'intelligibilité en fonction de la fréquence de coupure est en réalité une sigmoïde qui varie de 0 à 100% avec une partie linéaire comprise entre 353 et 2000 Hz. Sur cette partie l'intelligibilité progresse de 35% par octave (régression linéaire : $p < 10^{-3}$; $R^2 = 99\%$; $N = 20$). Cela est cohérent avec l'étude publiée par French et Steinberg (1947), travaux qui avaient été financés par une société de télécommunication, et qui indiquaient l'évolution de l'intelligibilité des syllabes en fonction de la fréquence de coupure passe-bas par une sigmoïde.

Le modèle linéaire sur la zone 353-2000 Hz, nous permet d'établir une base de référence pour situer les résultats des malentendants par rapport à ceux des normo-entendants. Nos témoins ont été testés dans l'oreille droite uniquement et sans entraînement au test ; les résultats que nous avons obtenus sont probablement sous estimés dans la mesure où il n'y a pas eu d'apprentissage des listes. Il se peut également que les résultats soient légèrement différents avec un test sur l'oreille gauche.

4.2. Apport des appareils

Les deux systèmes testés nécessitent un gros travail de rééducation auditive, particulièrement l'implant EAS. Les performances obtenues avec le Naïda apparaissent prometteuses, même si individuellement, un patient (patient 1) comprend moins bien oreilles appareillées que oreilles nues. Pour ce patient, il s'agit d'un premier appareillage, sa surdité date d'une quinzaine d'année et c'est celui qui éprouve le plus de difficultés à comprendre avec son appareil. Cette observation laisse à penser que les patients qui ont eu de longues durées de surdité sont probablement habitués à n'utiliser que leurs graves pour comprendre. Pour ces patients qui sont habitués à ne pas utiliser les hautes fréquences, il sera sans doute difficile d'exploiter les sons aigus si on tente de les restituer. Le patient Naïda n°2 est celui qui a la surdité la plus ancienne mais ses aigus ont pratiquement toujours été stimulés avec ses précédents appareillages, on peut donc penser qu'il accepte mieux la restitution des hautes fréquences malgré la longue durée de privation auditive. Quant aux patients n° 3 et 4, nous ne connaissons pas leur durée de privation, mais ce sont des patients qui étaient déjà appareillés auparavant depuis plusieurs années et qui semblent avoir le même profil que le patient n°2. L'avantage incontournable de la compression fréquentielle reste toutefois la réduction du larsen, du fait que le message fréquentiel de sortie de l'appareil est différent du message d'entrée. Cela nous permet de délivrer des puissances beaucoup plus importantes, là où les appareils à amplification conventionnelle atteignent leur limite de gain critique, ce qui est souvent le cas pour le type de surdité que nous avons appareillé. L'inconvénient de ces systèmes sera sans doute la difficulté à renouveler l'appareillage. Lorsque l'appareil arrivera en fin de vie, le malentendant aura été habitué pendant les 5 à 8 années de port de l'appareil, à entendre un son qui présentait en quelque sorte des « distorsions », et dans l'idéal il faudra que l'appareil qui succède présente les mêmes distorsions pour que le patient accepte le nouveau son. Une ouverture de cette étude serait de comparer les situations « amplification sans compression » versus « amplification avec compression », ce qui permettrait de confirmer le bénéfice de la compression fréquentielle.

Globalement, les résultats avec l'implant EAS sont plutôt positifs. Pour le patient 2, la compréhension à J3 mois ne dépasse pas les 30%, mais pour les 3 autres patients implantés EAS, les résultats sont déjà très encourageants. Les patients implantés EAS 1 et 2 sont ceux qui ont les plus longues durées de privation auditive (18 et 19 ans de privation) et ce sont ceux qui ont les moins bonnes performances à l'audiométrie vocale avec l'appareillage. Tout comme le patient Naïda n°1, ces patients qui ont une longue durée de

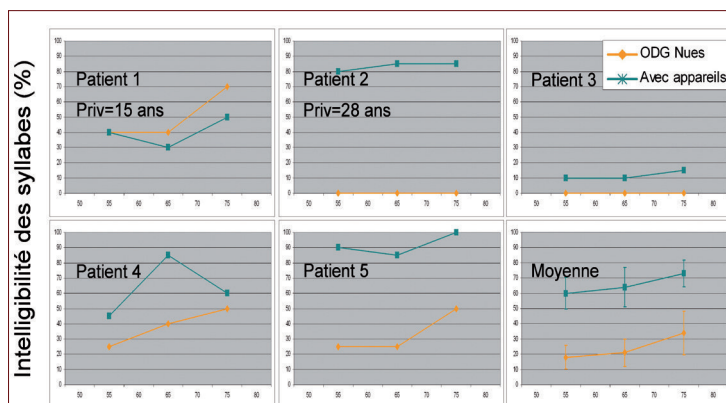


Figure 12. Audiométrie vocale en champ libre, oreilles nues et oreilles appareillées avec un Naïda. Le graphe en bas à droite représente la moyenne des audiométries vocales des 5 patients testés, la barre d'erreur correspond à l'erreur standard, les patients comprennent significativement mieux avec les appareils (test t par population, $p < 0.05$). La durée de privation de l'oreille (ou des oreilles) appareillée(s) est indiquée lorsque nous la connaissons.

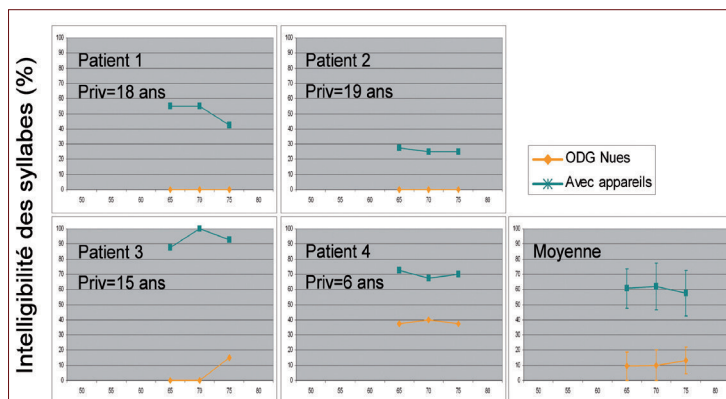


Figure 13. Audiométrie vocale en champ libre, oreilles nues et oreilles appareillées avec un implant EAS. Le graphe en bas à droite représente la moyenne des audiométries vocales des 4 patients testés, la barre d'erreur correspond à l'erreur standard, les patients comprennent significativement mieux avec leur implant (test t par population, $p < 0.05$). Pour chaque patient, nous indiquons la durée de privation de l'oreille appareillée.

surdité, acceptent difficilement la restitution des informations haute fréquence par l'implant.

Les patients implantés ont de meilleurs seuils en audiométrie tonale avec le système EAS et ils ont accès à plus d'information dans les hautes fréquences que les patients appareillés en Naïda, mais leur compréhension est moins bonne. Nous pouvons l'expliquer d'une part parce que les patients implantés EAS ont des surdités plus importantes au départ et d'autre part parce qu'il faut sans doute une rééducation plus longue pour intégrer la sensation auditive fournie par la stimulation bimodale.

La variabilité des résultats avec ces systèmes est sans doute la conséquence de critères de sélection inadaptés ou mal définis avec les moyens disponibles. Il est possible que les tests actuels d'audiométrie tonale et vocale atteignent leurs limites. Si nous regardons les résultats du test de vocale filtrée, tous les malentendants testés dans le protocole d'évaluation ont une courbe qui sature en intelligibilité lorsque la fréquence de coupure augmente. Cela signifie

qu'aucun n'utilise ses hautes fréquences pour comprendre. A priori nous pensons que les hautes fréquences peuvent leur être restituées soit par la partie implant cochléaire de l'EAS, soit par la compression fréquentielle. Toutefois, si la surdité est trop ancienne, ces systèmes risquent de ne pas être aussi efficaces qu'on pourrait l'espérer, les patients étant habitués depuis plusieurs années à ne comprendre qu'avec leurs graves. Les résultats du test de vocale filtrée montrent même que certains patients comprennent parfois mieux la parole filtrée que non filtrée (patient Naïda 1 OD, patient Naïda 4 OD, patient EAS 2 OD et patient EAS 4 OD). Dans ce cas de figure, l'audiométrie vocale classique sous-évalue les performances du patient, c'est donc qu'elle n'explore pas suffisamment ses capacités à utiliser les basses fréquences. Le test d'audiométrie vocale filtrée que nous avons utilisé permet ainsi de mieux explorer la contribution des basses fréquences chez le malentendant et de voir si il utilise encore ses hautes fréquences pour comprendre.

5

Conclusion

Le test d'audiométrie vocale filtrée que nous avons mis en oeuvre peut être un outil utile pour explorer de façon plus précise l'utilisation ou non des hautes fréquences chez le malentendant. Il pourra être utilisé par la suite dans le but de réaliser un bilan pronostique afin de déterminer « quel système est le mieux adapté à quel profil de patient ». Cependant, nos résultats montrent que la réussite de l'appareillage ne dépend pas seulement de la technologie utilisée, mais elle dépend sans doute également de facteurs propres au patient, en particulier la durée de privation auditive.

Dans notre étude, le Naïda a été évalué sur 5 patients et l'implant EAS a été évalué sur 4 patients, notre effectif est faible pour pouvoir tirer des conclusions définitives avec nos résultats. Un échantillon plus grand et une analyse de l'effet de la durée de privation permettraient de confirmer si, effectivement, l'ancienneté de la surdité modifie les performances avec l'EAS et avec la compression fréquentielle.

Dans notre étude, nous avons constaté que les patients comprenaient en général mieux avec le Naïda ; il faudrait maintenant quantifier la part d'amélioration due à l'amplification et celle due à la compression de fréquences, pour être moins dépendant de l'échantillon.

Les aides auditives à compression non-linéaire de fréquence et l'implant électrique-acoustique ont un avenir très prometteur pour l'appareillage des surdités supérieures à 70 dB au-delà de 1000 Hz. Pour les utiliser de manière optimale, il est nécessaire de bien définir les indications d'appareillage afin d'exploiter leur potentiel au maximum.

6

Bibliographie

Adunka O. and Kiefer J., "Impact of electrode insertion depth on intra-cochlear trauma", *Otolaryngol. Head Neck Surg.*, 135(3) (Sep 2006), pp. 374-382.

Adunka O., Unkelbach M.H., Mack M.G., Radeloff A., Gstoettner W., "Predicting basal cochlear length for electric-acoustic stimulation", *Arch. Otolaryngol. Head Neck Surg.*, vol. 131 (June 2005), pp. 488-492.

Baskent D. and Shannon R.V., "Combined effects of frequency compression-expansion and shift on speech recognition", *Ear Hear.*, 28(3) (2007), pp. 277-289.

Briggs R.J.S., Tykocinski M., Xu J., Risi F., Svehl M., Cowan R., Stöver T., Erfurt P., Lenarz T., "Comparison of Round Window and Cochleos-tomy Approaches with a Prototype Hearing Preservation Electrode". *Audiol. Neurotol.*, vol. 11(suppl1) (2006), pp. 42-48.

Ching T.Y.C., Dillon. H., Byrne D., "Speech recognition of hearing-impaired listeners: predictions from audibility and the limited role of high-frequency amplification", *J. Acous. Soc. Am.*, 103(2) (Feb 1998), pp. 1128-1140.

Cutler A., Mehler J., Norris D., Segui J., "A language-specific comprehension strategy", *Nature*, 304(5922) (1983 Jul 14-20), pp. 159-160.

Cutler A., Mehler J., Norris D., Segui J., "Limits on bilingualism", *Nature*, 340(6230) (1989 Jul 20), pp.229-230.

Dorman M.F., Spahr A.J., Loizou P.L., Dana C.J., Schmidt J.S., "Acoustic simulations of combined electric and acoustic hearing (EAS)", *Ear. Hear.*, 26(4) (Aug 2005), pp. 371-380.

French N.R., Steinberg J.C., "Factors governing the intelligibility of speech sounds", *J. Acous. Soc. Am.*, 19(1) (Jan 1947), pp. 90-119.

Gantz B.J. and Turner C.W., "Combining acoustic and electrical hearing", *Laryngoscope*, vol. 113 (Oct 2003), pp. 1726-1730.

Gantz B.J., Turner C., Gfeller K.E., Lowder M.W., "Preservation of hearing in cochlear implant surgery : advantages of combined electrical and acoustical speech processing", *Laryngoscope*, 115(5) (May 2005), pp. 796-802.

Gifford R.H., Dorman M., Spahr A.J., McKarns S.A., "Effect of digital frequency compression (DFC) on speech recognition in candidates for combined electric and acoustic stimulation (EAS)", *J. Speech Lang. Hear. Res.*, 50(5) (2007), pp. 1194-1202.

Gilbert G., "Fonctions d'importance fréquentielle pour la reconnaissance de la parole : application et amélioration d'une approche corrélationnelle", thèse présentée le 15/12/2003, Université Claude BERNARD – LYON 1.

Gilbert G., Micheyl C., "Influence of competing multi-talker babble on frequency-importance functions for speech measured using a correlational approach", *Acta Acustica united with Acustica*, 91 (2005), pp. 145-154.

Gstoettner W.K., Helbig S., Maier N., Kiefer J., Radeloff A., Adunka O.F., "Ipsilateral electric acoustic stimulation of the auditory system: results of long-term hearing preservation", *Audiol. Neurotol.*, vol. 11(suppl1) (2006), pp. 49-56.

Gstoettner W., Van De Heyning P., Fitzgerald O'Connor A., Morera C., Sainz M., Vermeire K., McDonald S., Cavallé L., Helbig S., Valdecasas J.G., Anderson I., Adunka O., "Electric acoustic stimulation of the auditory system : results of a multi-centre investigation", *Acta Otolaryngol.*, 128(9) (Sep 2008), pp. 968-975.

Gstoettner W., Kiefer J., Baumgartner W.D., Pok S., Peters S., Adunka O., "Hearing preservation in cochlear implantation for electric acoustic stimulation", *Acta Otolaryngol.*, vol. 124 (2004), pp. 348-352.

Von Ilberg C., Kiefer J., Tillein J., Pfenningdorff T., Hartmann R., Stürzebecher E., Klinke R., "Electric-Acoustic Stimulation of the Auditory System" *ORL*, vol. 61 (1999), pp. 334-340.

James C., Fraysse B., Deguine O., Lenarz T., Mawman D., Ramos A., Ramsden R., Sterkers O., "Combined Electroacoustic Stimulation in Conventional Candidates for Cochlear Implantation", *Audiol. Neurotol.*, vol. 11(suppl1) (2006), pp. 57-62.

Jose H.A.L., Asano F., Suzuki Y., Sone T., "Adaptive feedback cancellation with frequency compression hearing aids", *J. Acous. Soc. Am.*, 94(6) (Dec 1993), pp. 3248-3254.

Kiefer J., Böhnke F., Adunka O., Arnold W., "Representation of acoustic signals in the human cochlea in presence of a cochlear implant electrode", *Hear. Res.*, 221(1-2) (Nov 2006), pp. 36-43.

Kiefer J., Gstoettner W., Baumgartner W., Pok S.M., Tillein J., Ye Q., Von Ilberg C., "Conservation of low-frequency hearing in cochlear implantation," *Acta Otolaryngol.*, vol. 124 (2004), pp. 272-280.

Kiefer J., Pok M., Adunka O., Stürzebecher E., Baumgartner W., Schmidt M., Tillein J., Ye Q., Gstoettner W., "Combined electric and acoustic stimulation of the auditory system : results of a clinical study", *Audiol. Neurotol.*, 10(3) (May-Jun 2005), pp. 134-144.

Lenarz T., Stover T., Buechner A., Paasche G., Briggs R., Risi F., Pesch J., Battmer R.D., "Temporal Bone Results and Hearing Preservation with a New Straight Electrode", *Audiol. Neurotol.*, vol. 11(suppl1) (2006), pp. 34-41.

McDermott H.J., Lech M., Kornblum M.S., Irvine D.R.F., "Loudness perception and frequency discrimination in subjects with steeply sloping hearing loss: possible correlates of neural plasticity", *J. Acous. Soc. Am.*, 104(4) (Oct 1998), pp. 2314-2325.

Reed C.M., Hicks B.L., Braida L.D., Durlach N.I., "Discrimination of speech processed by low-pass filtering and pitch-invariant frequency lowering", *J. Acous. Soc. Am.*, 74(2) (August 1983), pp. 409-419.

Simpson A., "Frequency-lowering devices for managing high-frequency hearing loss: a review", *Trends Amplif.*, 13(2) (June 2009), pp. 87-106.

Skarzynski H., Lorens A., D'Haese P., Walkowiak A., Piotrowska A., Sliwa L., Anderson I., "Preservation of residual hearing in children and post-lingually deafened adults after cochlear implantation : an initial study", *ORL* 64(4) (Jul Aug 2002), pp. 247-253.

Skarzynski H., Lorens A., Piotrowska A., Anderson I., "Partial deafness

cochlear implantation provides benefit to a new population of individuals with hearing loss," *Acta Otolaryngol.*, vol. 126 (2006), pp.934-940.

Thai-Van H., Michey C., Moore C.J., Collet L., "Enhanced frequency discrimination near the hearing loss cut-off: a consequence of central auditory plasticity induced by cochlear damage?", *Brain*, 126 (2003), pp.2235-2245.

Thai-Van H., Michey C., Norena A., Collet L., "Local improvement in auditory frequency discrimination is associated with hearing-loss slope in subjects with cochlear damage", *Brain*, 125 (2002), pp.524-537.

Turner C.W., "Hearing loss and the limits of amplification", *Audiol. Neurotol.*, 11(suppl1) (2006), pp. 2-5.

Turner C.W., Gantz B.J., Vidal C., Behrens A., Henry B.A., "Speech recognition in noise for cochlear implant listeners : benefits of residual acoustic hearing", *J. Acous. Soc. Am.*, 115(4) (Apr 2004), pp. 1729-1735.

Turner C.W. and Hurtig R.R., "Proportional frequency compression of speech for listeners with sensorineural hearing loss", *J. Acous. Soc. Am.*, 106(2) (August 1999), pp. 877-886.

Vickers D.A., Moore B.C.J., Baer T., "Effect of low-pass filtering on the intelligibility of speech in quiet for people with and without dead regions at high frequencies", *J. Acous. Soc. Am.*, 110(2) (Aug 2001), pp. 1164-1175.

Xiao X., Hu G., Liu C., Liu J., "Improving speech intelligibility of severely hearing-impaired people by frequency-lowering technique", *Scientific Research and Essay*, 3(6) (June 2008), pp. 259-266.

I.11. Variabilité des performances en EAS

La littérature sur l'EAS montre que certains candidats ont de bonnes performances avec l'EAS. Cela signifie que la méthode fonctionne et qu'elle est validée. Mais nous constatons également des échecs et tous ne sont pas expliqués. Il y a des inconnues sur les résultats de l'EAS et c'est ce que nous essayons de mieux appréhender. Une partie des échecs avec l'EAS sont dus au fait que l'audition résiduelle a été détruite, tandis que pour d'autres cas l'EAS ne fournit aucun bénéfice par rapport à la stimulation électrique seule, bien que l'audition résiduelle soit partiellement préservée après la chirurgie.

Dans une étude de Gstoeckner et al. (2008), 12 sujets sur 18 avaient suffisamment d'audition préservée après la chirurgie pour bénéficier de l'EAS, pourtant sur ces 12 sujets 6 ont choisi de ne pas utiliser le module acoustique, car ils sentaient que la qualité du son était dégradée, que l'amplification était insuffisante, ou qu'ils n'appréciaient pas la stimulation acoustique combinée à l'implant (c'est probablement ce qui se produit chez le 1^{er} patient que nous présentons dans notre étude de cas, plus tard dans ce manuscrit). Ce constat remet alors en question les résultats de Von Ilberg et al. (1999) lorsqu'ils affirment que la combinaison d'une stimulation acoustique et électrique n'engendre pas d'interférences perturbantes et que le système auditif central est capable d'intégrer cette stimulation bimodale.

Chez Kiefer et al. (2005), sur 13 sujets implantés, 11 ont pu bénéficier de l'EAS et chez 2 sujets EAS sur 11, à J1an, le score de reconnaissance des mots monosyllabiques est encore inférieur ou égal à 40%.

La même année, Gantz et al. (2005) trouvent une amélioration importante de la reconnaissance des mots monosyllabiques chez 10 sujets sur 11 qui ont eu leur audition résiduelle préservée, pour la condition EAS versus électrique seule. Pour le 11^{ème} sujet qui n'a une amélioration que très légère, son intelligibilité des mots monosyllabiques est passée de 20% pré-implantation à environ 25% avec l'implant seul et 28% avec l'EAS, l'EAS ici n'est donc pas aussi efficace qu'on peut l'espérer. De même, dans leur étude de 2009 Gantz *et al.* ont trouvé que chez les patients EAS qui avaient de mauvaises performances, 91% de la variance pouvait être expliquée par le score de reconnaissance de la parole pré-implantation et par la durée de surdité.

Dans les études où les participants sont implantés cochléaires et utilisent une aide auditive du côté opposé à l'implant, des résultats analogues sont constatés. Par exemple dans une étude de Dunn *et al.* (2005), sur 12 sujets testés seulement 4 tirent un bénéfice significatif de l'EAS pour la reconnaissance des mots CNC dans le silence. Pour Mok *et al.* (2006) seulement 2 sujets sur 14 testés montrent un bénéfice significatif de l'EAS pour des phonèmes CNC et certains de leurs sujets ont même des performances moins bonnes en condition EAS par rapport à la condition implant seul. En 2007, Novak *et al.* présentaient les résultats obtenus chez 9 patients adultes avec une perte auditive en pente abrupte sur les hautes fréquences, qui utilisent un implant cochléaire et une aide auditive contralatérale. Les sept sujets qui avaient une perte progressive avec une durée de surdité variant de 1 à 20 ans ont vu leur compréhension de la parole s'améliorer significativement avec la stimulation bimodale. En revanche les deux patients qui avaient une surdité congénitale avec une durée de privation de 37 et 44 ans n'ont montré qu'une légère, voire pas d'amélioration. Les auteurs expliquent ces mauvaises performances par une moins bonne survie neuronale et une durée de privation auditive plus longue chez les deux sourds congénitaux. Pour Kong et Braida (2011), ce manque de bénéfice résulterait d'un problème d'incompatibilité des 2 appareils (ou modalités) étant donné qu'une oreille reçoit une stimulation acoustique et que l'autre oreille reçoit une stimulation électrique.

Dans l'étude de Kong et Braida (2011), les sujets ont toujours reçu des entrées sensorielles de parole dégradée depuis l'enfance. De ce fait leur représentation interne des phonèmes est essentiellement basée sur une perception basse fréquence. Dès l'instant où les sujets reçoivent un implant, ils ne sont plus capables d'identifier le phonème car en stimulation bimodale le phonème ne correspond plus à la représentation interne qu'ils en ont. Ces patients ont donc une capacité réduite pour intégrer les indices de parole haute fréquence nouvellement perçus et pour identifier les stimuli en condition bimodale. Ces phénomènes d'imperfections des représentations internes du stimulus se retrouvent également chez les sourds post-linguaux qui ont une surdité ancienne (Sagi *et al.*, 2010).

I.12. Bilan sur les performances avec les différents systèmes

Quel que soit le type d'appareillage, les résultats montrent toujours de très larges disparités interindividuelles. Ce constat est probablement le résultat d'indications mal adaptées pour les différentes technologies. Dans tous les travaux décrits dans cette introduction, nous avons vu que les performances étaient souvent liées à des facteurs propres aux patients tels que la durée de la surdité ou le degré d'audition résiduelle.

Chacun des systèmes va nécessiter un certain degré de « plasticité cérébrale ». Dans le cas de l'aide auditive conventionnelle, le sujet va devoir apprendre à ne comprendre qu'avec ses restes auditifs, simplement avec la portion fréquentielle encore exploitable. Dans le cas d'appareils à compression fréquentielle, le sujet utilise son audition résiduelle mais il apprend au fil du temps à utiliser des sons auxquels il n'avait pas accès auparavant pour comprendre.

Pour l'implant cochléaire, le problème est beaucoup plus long et complexe puisque cela impose une réattribution fréquentielle. L'éducation ou la rééducation auditive est donc plus facile chez les cerveaux jeunes ou en cours de développement et chez les sujets qui ont des durées de privation plus courtes.

Dans le cas de la stimulation bimodale, la rééducation est probablement plus difficile dans la mesure où le patient doit non seulement apprendre à comprendre avec l'implant cochléaire (réattribution fréquentielle), mais il doit surtout apprendre à combiner des informations provenant de deux modalités différentes mais qui peuvent être redondantes. Le travail risque d'être doublement difficile pour ces patients, mais en contrepartie, lorsque les performances seront optimales, ces patients auront probablement accès à beaucoup plus d'indices acoustiques que les autres populations de patients.

Problématiques émergentes pour l'EAS

***Indications**

L'EAS est efficace chez la majorité des patients et le bénéfice fourni par l'EAS par rapport à l'IC n'est pas discutable, en particulier dans les conditions d'écoute difficiles, où le bruit masque la parole (Olson et Shinn, 2008). Toutefois, si nous nous référons aux publications qui présentent les résultats individuels des sujets testés, nous constatons qu'il y a certains patients pour qui la stimulation bimodale ne fournit aucun bénéfice par rapport à leur aide auditive seule ou à leur implant seul, et dans certains cas les auteurs observent une baisse des performances lorsque les deux systèmes sont combinés (Hamzavi et al., 2004 ; Dunn et al., 2005 ; Luntz et al., 2005 ; Mok et al., 2006). Certains supposent que cette variabilité pourrait être partiellement expliquée par des informations conflictuelles provenant de l'aide auditive et de l'implant et des interférences bimodales (Mok et al., 2006). Il est donc nécessaire de comprendre et d'expliquer cette disparité des performances chez les patients implantés EAS. En particulier, il nous semble important 1/ de mieux cibler les indications de l'EAS ; 2/ d'évaluer dans quelle mesure les facteurs « durée de privation » et « âge d'apparition de la surdité » vont conditionner les performances du sourd partiel.

***Stratégies de codage**

Les portes-électrodes conçus pour l'EAS sont généralement plus courts que les porte-électrodes standard. Par exemple la firme Cochlear proposent deux porte-électrodes pour l'EAS. Le porte-électrode court Hybrid-S mesure 10 mm et possède 6 contacts (Gantz et Turner, 2003 ; Patrick et al., 2006), mais ce faisceau a l'inconvénient d'être trop court et risque de ne pas fournir des performances optimales si le sujet implanté perd son audition résiduelle post-implantation. Pour remédier à cela, Cochlear propose un porte-électrode plus long Hybrid-L (ou Hybrid-L24) qui mesure 16 mm (longueur active de 15mm) et possède 22 canaux de stimulation (Lenarz et al., 2006 ; Lenarz et al., 2009 ; Driscoll et al., 2011). Pour avoir une référence, le porte-électrode standard mesure 25mm et possède 22 canaux sur une longueur active de 15.5 mm (Zeng et al., 2008).

Pour la firme Med-El, le porte-électrode Flex^{EAS} conçu pour les chirurgies EAS mesure 24 mm, il possède 12 canaux (7 paires de contacts sur les électrodes basales et 5 contacts uniques sur les électrodes apicales) sur 20,9 mm avec un espace inter-électrode de 1,9 mm. Afin d'augmenter la flexibilité, ce porte-électrode possède 7 paires de contacts sur les électrodes basales et 5 contacts uniques sur les électrodes apicales (Adunka et al., 2004). Pour une comparaison, le porte-électrode standard mesure 31,5 mm, il possède 12 canaux sur 26,4 mm avec un espace inter-électrode de 2,4 mm, toutes les électrodes possèdent 2 contacts. L'espace inter-électrode est donc réduit de 20% sur le porte-électrode Flex^{EAS}. Actuellement, Med-El est en train de développer l'électrode Flex 28 (28 mm) qui sera un intermédiaire entre la Flex Soft et la Flex^{EAS}.

Chez les 2 principales firmes qui développent l'EAS, les distances inter-électrodes sont réduites par rapport aux porte-électrodes standards, ce qui va probablement favoriser les interactions inter-canaux, phénomène qui est souvent responsable d'une dégradation de la perception de la parole (Fu et Shannon, 1999). Ce phénomène d'interaction est naturellement présent du fait que l'implant dans la rampe tympanique est entouré de périlymphe, conducteur d'électricité, d'où une diffusion du courant lorsque l'implant stimule.

Dans le cas de l'EAS (avec les implants Med-El qui font l'objet de nos travaux), le phénomène va être accentué pour plusieurs raisons :

- Le porte-électrode est inséré dans le tour basal de la cochlée, or les études histopathologiques (Spoendlin et Schrott, 1988; Nadol et al., 1989; Nadol et al., 2001) ont mis en évidence que cette section de la cochlée est celle qui a le plus faible capital de fibres auditives et de cellules ganglionnaires. Pour induire des perceptions auditives, il faut donc délivrer des stimulations plus importantes dans cette région. Par conséquent elle est plus sensible aux interactions par rapports aux autres segments de la cochlée.

- Le porte-électrode Flex^{EAS} possède moins de contacts sur les 5 électrodes apicales, les surfaces de stimulation sont donc réduites par rapport au porte-électrode standard. Cela va augmenter la densité de charge sur les contacts (McCreery et al., 1990; Shannon et al., 1992), ce qui va augmenter les interactions (Huang et Shepherd, 2000) et augmenter l'impédance (Xu et al., 1993). Et pour induire une sensation auditive, il faudra probablement une plus forte stimulation (Xu et al., 1993).

- Pour limiter le risque d'endommager les structures fragiles de la cochlée, des études récentes ont montré que des insertions directes dans la fenêtre ronde ou proches de la fenêtre ronde étaient favorables par rapport à des insertions par cochléostomie (Adunka et al., 2004 ; Briggs et al., 2005). Avec une insertion par la fenêtre ronde, le porte-électrode est plus proche du mur latéral de la cochlée, les contacts vont être éloignés des corps cellulaires et il faut un courant de stimulation plus important pour obtenir une perception auditive (Shepherd et al., 1993; Cohen et al., 2006; Spelman, 2006). Cela résulte en une plus large diffusion du courant d'excitation, d'où une augmentation des interactions (Xi et al., 2009; Hughes et Abbas, 2006).

Actuellement il n'existe pas de stratégie de codage spécifique à l'EAS et les fabricants d'implants se contentent d'utiliser les stratégies de l'implant qu'ils appliquent à l'EAS. Constatant que les porte-électrodes pensés pour l'EAS favorisent malheureusement les interactions, nous pouvons nous demander si les stratégies dédiées à l'implant cochléaires sont vraiment optimales pour l'EAS. Une stratégie visant à limiter les interactions type N-of-M ou à cadence réduite pourrait probablement être plus adaptée à cette population de patients. Récemment, Wilson (2011) a déposé un brevet dans lequel il recommande pour l'EAS, d'utiliser des cadences de stimulation très lentes (< 100 Hz) pour les canaux hautes fréquences. Ce document est sans doute une première étape pour l'évolution des stratégies de codage de l'EAS.

Travaux réalisés et objectifs

Nos travaux de recherche avaient 2 objectifs. Le premier objectif était d'identifier les causes d'échec de la stimulation bimodale. Le second objectif était de proposer des solutions en vue d'améliorer le traitement du son pour l'EAS. Aujourd'hui le signal délivré par l'EAS est calqué sur le traitement de l'implant cochléaire (répartition fréquentielle, cadence de stimulation, stratégie de codage). Nous savons que cela fonctionne mais nous ne savons pas si cela est optimal pour une stimulation bimodale et à ce jour peu de personnes se sont interrogées à ce sujet.

Nous avons dans un premier temps réalisé une étude préliminaire chez le normo-entendant, pour mettre en place un modèle d'intelligibilité en fonction de la bande passante, pour de la parole filtrée passe-bas. Ce travail nous a permis de recueillir des valeurs normatives que nous avons utilisées dans la suite de nos recherches.

Dans une seconde étude, nous avons évalué l'utilisation des basses fréquences chez des sujets (non implantés) ayant une surdité partielle. Pour cela nous avons mis en place un test d'audiométrie vocale filtrée passe-bas. Les résultats obtenus ont mis en évidence une grande disparité des performances chez ces sujets non implantés, ce qui pourrait expliquer que la disparité des performances se retrouve post-implantation EAS. Tous ces sujets n'exploitent pas de la même façon leurs basses fréquences résiduelles pour comprendre. Nous avons cherché à évaluer l'impact de l'âge d'apparition de la surdité et de la durée de la surdité sur les capacités à utiliser les sons graves. Il apparaît que ces facteurs sont significativement corrélés aux performances ; il est donc important de les prendre en considération lorsque le clinicien propose de l'EAS à un malentendant sourd partiel.

Dans une troisième étude, nous avons proposé ce même test d'audiométrie vocale filtrée passe-bas en vue d'évaluer si un sujet atteint de surdité partielle pouvait être un bon candidat à l'EAS. Le but ici était de mieux catégoriser les différents profils de patients et de mieux cibler les indications de l'EAS. Le principe de l'EAS étant de placer un faisceau d'électrodes dans le tour basal de la cochlée, nous sommes quasiment certains que les cellules hautes-fréquences ne seront plus fonctionnelles post-implantation. Avant d'envisager l'EAS et donc de condamner les hautes fréquences, il est indispensable de s'assurer que le patient ne les utilise pas acoustiquement pour comprendre. Le test d'audiométrie filtrée va ainsi mettre en évidence l'utilisation ou non des hautes-fréquences pour la compréhension. S'il apparaît que le patient ne les utilise pas et qu'il est dans les indications audiométriques pour l'EAS, la stimulation bimodale pourra lui être proposée.

Le quatrième sujet présenté est une étude longitudinale menée chez 8 patients implantés EAS à Lyon. Nous avons recueilli et analysé les performances pour différentes mesures perceptives telles que l'audiométrie vocale filtrée passe-bas pré- et post-implantation, l'audiométrie en

condition acoustique, électrique et EAS, ainsi que des audiométries vocales dans le silence et dans le bruit lorsque ces données ont pu être mesurées chez les patients.

Dans une cinquième étude, nous avons réalisé des simulations d'implant électro-acoustiques, à l'aide d'un vocoder hybride, pour différentes configurations d'audition résiduelle et différents nombres de canaux sur la partie implant simulé. Des sujets normo-entendants écoutaient ces simulations d'EAS et répétaient ce qu'ils avaient compris. Les tests ont été réalisés dans le silence et à différents rapports signal/bruit. L'objectif était d'évaluer l'intelligibilité transmise par la partie acoustique, par la partie électrique et par les deux parties combinées. La finalité est de pouvoir prédire le nombre de canaux nécessaires pour obtenir une bonne intelligibilité. La perspective étant de réduire le nombre de canaux sur le porte-électrode de l'implant pour réduire les interactions, sans pour autant dégrader la compréhension.

Pour notre sixième étude, nous sommes partis d'un article de Kong et Carlyon (2007), qui justifiaient le bénéfice de la stimulation bimodale par un effet « superadditif » de la combinaison des informations acoustiques et électriques. Par des simulations en utilisant le même protocole que pour l'étude précédente, nous avons recueilli les scores pour les conditions « stimulation acoustique seule », « stimulation électrique seule » et « EAS ». En appliquant différents modèles d'intégration que nous avons trouvés dans la littérature, nous avons mis en évidence que le modèle utilisé par Kong et Carlyon (2007) n'était pas optimal et que les performances en condition EAS pouvaient être expliquées par des modèles d'intégration mieux adaptés.

Enfin, dans une septième étude, nous utilisons un autre vocoder hybride pour évaluer si une stratégie de codage à extraction de pics ou à cadence lente pourrait être plus adaptée pour l'EAS. Actuellement, les porte-électrodes conçus pour l'EAS ont des contacts plus proches les uns des autres et des surfaces d'électrodes plus petites sur la partie apicale, en ce sens ils sont probablement plus sensibles aux interactions de canaux. Ceci pourrait expliquer des mauvaises performances chez certains patients EAS. Nous cherchons à évaluer, par des simulations acoustiques, si la stratégie de codage du son par l'implant a une influence sur les

performances de compréhension. Si nos résultats montrent que la stratégie utilisée n'a pas d'effet sur les performances, il pourrait être intéressant pour l'EAS d'utiliser une stratégie de codage qui limite ces phénomènes d'interactions. Ceci aurait pour conséquence une meilleure qualité du message électrique délivré au nerf auditif par l'implant et ainsi limiter les interactions.

II. Productions scientifiques

II.1. Mise au point d'un test de parole filtrée pour mieux évaluer les indications de l'implantation EAS

(article publié : Seldran F, Thai-Van H, Truy E, Berger-Vachon C, Collet L, Gallego S, Beliaeff M. A filtered speech test to better evaluate electric acoustic stimulation (EAS) candidacy. Cochlear Implants Int. 2010;11 Suppl 1:130-3)

La première partie de nos travaux a consisté à rechercher comment améliorer les tests pré-implantatoires afin d'affiner les indications de l'EAS. Nous sommes partis du constat que l'EAS est efficace chez la majorité des patients, mais qu'il existe certains patients pour qui l'EAS n'améliore pas les performances par rapport aux conditions « aide auditive seule » ou « implant seul ». Nous avons supposé que ceci était le résultat de critères d'implantations mal adaptés ou mal définis et il nous semblait important de développer un outil pour permettre de mieux sélectionner les patients à implanter. Actuellement, d'après les constructeurs et un consensus, les indications de l'EAS en audiométrie tonale sont une audition résiduelle significative avec des seuils de perception meilleurs que 65 dB HL jusqu'à 700-1000 Hz et une surdité sévère à totale au-delà de 1000 Hz. Quant à l'audiométrie vocale pré-implantation, elle doit être aux environs de 60% (Kiefer et al., 2002 ; Van de Heyning, 2010; Arnoldner et al, 2010).

La problématique de l'EAS vient du fait que l'insertion du porte-électrode sur la partie basale de la cochlée risque de détruire les cellules hautes fréquences de la cochlée. Avant de condamner ces cellules, il est important pour le clinicien de s'assurer qu'elles n'apportent rien lorsqu'elles sont stimulées acoustiquement. C'est pourquoi nous avons proposé de développer un outil qui permet d'explorer l'utilisation des hautes fréquences par le sujet candidat à l'EAS. Pour cela nous avons mis en place un test simple-mais malheureusement long-d'audiométrie vocale filtrée passe-bas. Le principe de ce test étant de voir si une augmentation de la bande passante (i.e. de l'audibilité) augmente l'intelligibilité du sujet. Si, au-delà d'une certaine fréquence de coupure, l'augmentation de la bande passante n'améliore pas l'intelligibilité, cela signifie que le sujet n'utilise pas ses hautes fréquences pour comprendre.

Dans un premier temps, nous avons cherché à évaluer l'intelligibilité en fonction de la bande passante chez le normo-entendant afin d'obtenir des données normatives. Vingt sujets normo-entendants ont participé à notre étude. Grâce aux données recueillies, nous avons pu établir un modèle auquel nous comparons les patterns de réponse des sujets malentendants testés, ce qui nous permet de vérifier si effectivement le sujet sourd partiel utilise ou non ses hautes fréquences pour comprendre.

Après avoir testé nos premiers sujets malentendants, nous avons pu constater des patterns de réponse très différents. En fonction des réponses des patients, nous avons fait des pronostics sur l'appareillage qui nous semblait le plus adapté pour chaque type de patient.

Ces données et nos recommandations cliniques sont présentées dans l'article 6 ci-après. Bien entendu, nos pronostics décrits dans la discussion de cet article étaient purement spéculatifs mais cohérents à ce stade de l'étude et méritent d'être confirmés par des tests post-implantation. Toutefois le message important de cet article était de mettre en évidence qu'il existe plusieurs profils de patients. Au-delà de l'audiométrie tonale et vocale, il est important d'explorer l'audition résiduelle du sujet par des tests complémentaires, avant de proposer une implantation EAS.

A Filtered Speech Test to Better Evaluate Electric Acoustic Stimulation (EAS) Candidacy

FABIEN SELDRAN, HUNG THAI-VAN, ERIC TRUY,
CHRISTIAN BERGER-VACHON, LIONEL COLLET, and
STÉPHANE GALLEG0

*CNRS UMR 5020, Laboratoire Neurosciences Sensorielles, Comportement,
Cognition — Hôpital Edouard HERRIOT, Service d'audiologie et
d'explorations orofaciales, Lyon, France*

FABIEN SELDRAN and MICHEL BELIAEFF

VIBRANT MED-EL Hearing Technology GmbH, Sophia-Antipolis, France

Introduction

Electric Acoustic Stimulation (EAS) is a new treatment option for partial deafness. Most of these patients are unable to use their high frequencies acoustically. Turner (2006) showed that high-frequency amplification was efficient only if the hearing loss is less than 60–80 dB HL. He suggested use of a filtered speech test when the clinician doubts the effectiveness of amplification. Despite hearing preservation, some EAS patients may not have improvement of speech understanding with the combination of electric and acoustic stimulation compared to the implant alone. For patients with high-frequency hearing loss, this study evaluated the intelligibility provided by the residual low frequencies in order to predict if a patient is a good EAS candidate.

Methods

French Fournier lists of dissyllabic words were used for this study. They were low-pass filtered at several cutoff frequencies every half octave (2000, 1414, 1000, 707, 500, 353, 250 Hz). Speech audiometry tests were performed on 20 normal-hearing listeners with unfiltered and filtered lists, at intensities between 10 and 80 dB SPL at 10 dB steps. These data were used to plot a series of curves of 'speech intelligibility depending on stimulation level' for each low-pass cutoff frequency.

Results obtained at intensities of 40 and 50 dB SPL were used to establish a normalized model of 'speech intelligibility depending on low-pass cutoff frequency'. The subjects were tested through an earphone in the right ear without prior practice. Test material was presented and they were asked to repeat the word they had heard, even if this word had no meaning. The number of syllables repeated correctly was scored.

As each list contains 10 dissyllabic words, each syllable repeated correctly was equal to 5 per cent recognition. All lists were presented in random order to prevent learning of the test material by the subjects.

All filtering was performed with 'Cool Edit Pro' software, utilizing the 'FFT Filter' function which reduces all spectral coefficients of the attenuated band by 70 dB. One CD was made for each cutoff frequency. A CD player was linked to an audiometer to present the test material.

Individuals with high-frequency hearing loss were tested in a similar fashion. For these subjects, 2 lists of 10 words were administered. Each syllable correctly repeated was equal to 2.5 per cent recognition. Testing was performed through earphones, just as for the normal-hearing subjects. The two ears were tested separately at a comfortable level for unfiltered speech (between 80 and 100 dB SPL). This same comfort level was maintained for filtered speech material.

Results

For normal-hearing subjects, a series of curves of 'speech intelligibility depending on stimulation level' were obtained (Figure 1a). Each curve corresponded to one cutoff frequency. With louder intensity levels, intelligibility increased up to 40 dB SPL. Above this level, intelligibility reached a maximum and saturated for all cutoff frequencies. Results obtained at 40 and 50 dB SPL were used to calculate mean results for all cutoff frequencies and to establish our model of 'speech intelligibility depending on low-pass cutoff frequency' (Figure 1b).

Similar tests using low-pass filtered speech had previously been performed in English (French & Steinberg, 1947). It was necessary to normalize our model using French speech material. Linear regression analysis between the cutoff frequencies of 353 and 2000 Hz demonstrated an increase of intelligibility of 35 per cent per octave

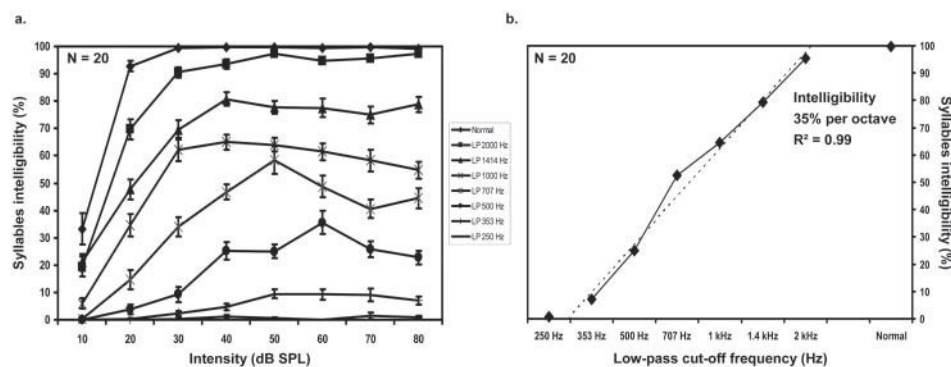


FIGURE 1a Percentage of syllables intelligible versus stimulation level for different low-pass cutoff frequencies. Mean values and standard errors obtained from 20 normal-hearing subjects.

FIGURE 1b Linear regression model of syllables intelligible (%) versus the logarithm of the low-pass cutoff frequency calculated from the mean results of 20 normal-hearing subjects at stimulation levels of 40 and 50 dB SPL. Slope is 0.35 per octave.

($N=6$, $R^2=0.99$, $p<0.0001$). Thus, greater high-frequency information improved speech understanding in normal hearing listeners.

Individuals with high-frequency hearing loss were tested. These subjects were classified into 4 categories depending on the contribution of the residual low-frequency hearing to speech intelligibility.

Figure 2 — example 1 depicts the first pattern of results. Performance was similar to normal-hearing listeners up to 700 Hz. Above 700 Hz, speech intelligibility remained stable and did not further improve. High-frequency information above 700 Hz was not useful to the patient.

Figure 2 — example 2 demonstrates the second pattern observed. Performance was similar to normal-hearing listeners up to 700 Hz. Beyond 700 Hz, speech intelligibility decreased as more high frequency information was provided. High-frequency information actually had a deleterious effect on speech intelligibility for this patient. This unusual pattern has been observed in previous studies (Vickers *et al.*, 2001; Hogan & Turner, 1998).

Figure 2 — example 3 illustrates the third pattern of results. For these individuals, additional high-frequency information improved speech recognition. However, performance was always inferior to normal-hearing subjects. For these subjects, high-frequency information was useful.

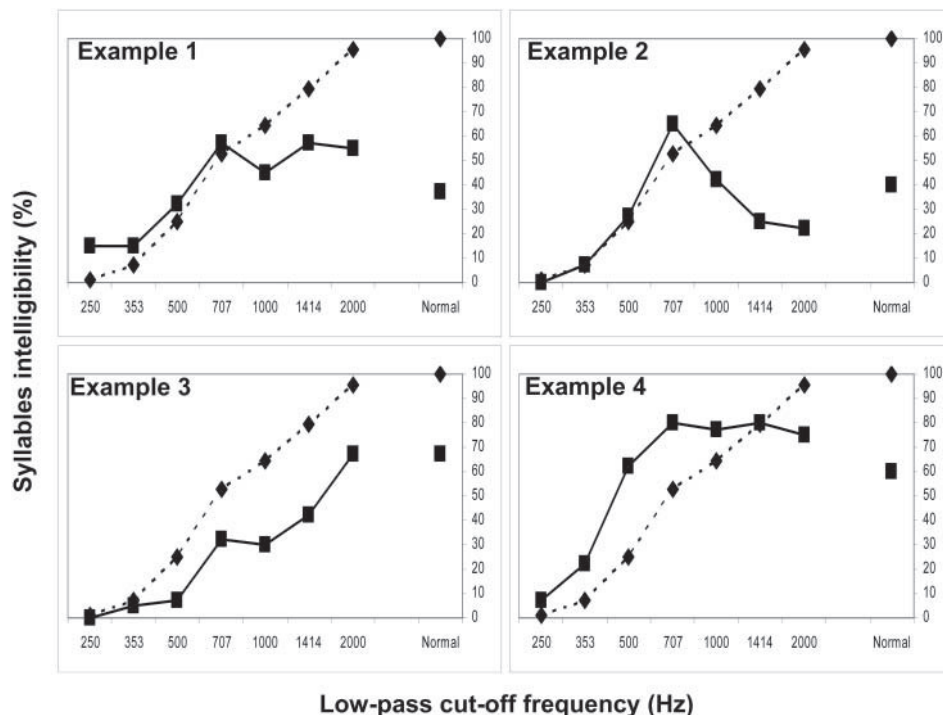


FIGURE 2 Four patterns of syllable intelligibility depending on the low-pass cutoff frequency for subjects with high-frequency hearing loss. For each pattern, the dashed line (♦) represents the model established from normal-hearing listeners and the solid line (■) represents the patient's result.

Figure 2 — example 4 shows the fourth pattern observed. This pattern is similar to the first pattern but speech understanding reached 80 per cent at 700 Hz. Likely, these patients have supra-normal abilities to use their low frequencies (Vestergaard, 2004).

Discussion

This study permitted categorization of individuals with high-frequency hearing loss according to their ability to use high-frequency information to improve speech recognition. According to these results, patients with patterns 1 and 2 did not use high-frequency information to understand speech. We assume these patients could benefit from electric stimulation in the high frequencies and would therefore be good EAS candidates.

Patients with pattern 3, despite their hearing loss, still use their high frequencies to understand speech. They might lose this ability with the insertion of an electrode array into the basal turn of the cochlea. It is therefore unknown if they are good EAS candidates.

Patients with pattern 4 have good speech understanding (80%) using only their low-frequency residual hearing. The risk of total deafness due to EAS surgery is too substantial to consider implantation in these individuals. Moreover, with such good performance using only their low-frequency residual hearing, it is not known if they will be able to use the high frequencies delivered by the implant.

Conclusion

The filtered speech test can determine the degree to which an individual uses their residual hearing to understand speech. This test may help the clinician to decide if EAS is the most appropriate solution. A larger study is planned in order to better define the indication criteria. In the future, when enough patients have undergone EAS implantation, we will be able to assess the prognostic value of the filtered speech test.

Bibliography

- French, N.R. & Steinberg, J.C. 1947. Factors Governing the Intelligibility of Speech Sounds. *J. Acous. Soc. Am.*, 19: 90–119.
- Hogan, C.A. & Turner, C.W. 1998. High-Frequency Audibility: Benefits for Hearing-Impaired Listeners. *J. Acous. Soc. Am.*, 104: 432–41.
- Turner, C.W. 2006. Hearing Loss and the Limits of Amplification. *Audiol. Neurotol.*, 11(Suppl 1): 2–5.
- Vestergaard, M.D. 2004. Benefit from Amplification of High Frequencies in Hearing Impaired: Aspects of Cochlear Dead Regions and Auditory Acclimatization. Technical University of Denmark, Ørsted-DTU Acoustic Technology, DK-2800 Lyngby.
- Vickers, D.A., Moore, B.C. & Baer, T. 2001. Effect of low-Pass Filtering on the Intelligibility of Speech in Quiet for People With and Without Dead Regions at High Frequencies. *J. Acous. Soc. Am.*, 110: 1164–75.

Correspondence to: fseldran@yahoo.fr

II.2. Effet de l'âge d'apparition de la surdité et de la durée de privation sur les performances de compréhension de la parole, dans le cas de surdité dans les hautes fréquences

(article publié : Seldran F, Gallego S, Micheyl C, Veuillet E, Truy E, Thai-Van H. Relationship between age of hearing-loss onset, hearing-loss duration, and speech recognition in individuals with severe-to-profound high-frequency hearing loss. J Assoc Res Otolaryngol. 2011 Aug;12(4):519-34)

Dans l'article 6 que nous venons de présenter, nous avons mis en évidence qu'à audiogramme équivalent tous les patients n'utilisaient pas les mêmes fréquences pour comprendre. Si nous prenons les exemples 3 et 4 de l'article précédent, ces deux patients ont la même intelligibilité en condition large bande (parole non filtrée), mais en condition filtrée leurs courbes de réponse sont très différentes. Le patient de l'exemple 4 va surexploiter ses 700 premiers Hz et ne pas utiliser les sons au-delà tandis que le patient de l'exemple 3 utilise toutes les fréquences mais il est sous-performant pour l'utilisation de tous ces sons. Ces différences ne peuvent pas être décelées sans utiliser un test d'audiométrie vocale filtrée passe-bas.

Il existe donc plusieurs profils de patients chez les sourds partiels et les tests réalisés sur une plus large population ont montré une très large disparité dans les performances et dans l'utilisation des basses fréquences pour la compréhension. Comment expliquer cette variabilité des performances ? La réponse la plus spontanée serait l'audibilité. Pourtant, les malentendants dont les données sont présentées dans l'article 6 ont été sélectionnés sur la base de leur audiogramme, à savoir que tous avaient une perte dans les hautes fréquences et étaient candidats à l'EAS d'après les critères actuels. Les patients testés ayant sensiblement les mêmes audiogrammes, cela signifie que les différences dans les profils sont probablement liées à des caractéristiques propres aux patients. C'est ce que nous avons tenté de comprendre dans cette seconde étude.

Dans l'introduction, nous avons vu que le facteur « durée de privation » avait un fort impact sur les performances du sujet implanté cochléaire ou EAS, à savoir que des longues durées de privation conduisaient à de mauvaises performances post-implantation. Pour les résultats avec l'implant cochléaire, il a été montré que la durée de privation pouvait expliquer jusqu'à 37% de la variance. C'est pourquoi il nous a semblé important d'évaluer dans quelle mesure ce paramètre pouvait également affecter les performances du sourd partiel pour l'utilisation de ses restes auditifs.

Notre hypothèse pour expliquer cela consiste à dire que les sourds partiels délaissent l'information haute fréquence au profit de l'information basse fréquence à laquelle ils ont encore accès acoustiquement et ce phénomène serait temps-dépendant. L'implant cochléaire est généralement introduit par la base de la cochlée, ce qui signifie qu'il stimule des cellules initialement vouées au traitement des informations hautes fréquences. Si le patient a effectivement développé une stratégie de compréhension basée sur la perception des sons graves au point qu'il ait délaissé ses hautes fréquences, il est concevable que les performances avec l'implant cochléaire ne soient pas optimales après plusieurs années de surdité.

Plusieurs auteurs ont déjà observé des résultats surprenants chez cette population atypique. Vestergaard (2003) a montré que les sourds comprenaient mieux la parole filtrée passe-bas par rapport à ce qui est prédit sur la base de l'audibilité. De même Moore et Vinay (2009) ont montré que les sujets qui présentaient des ZIC dans les hautes fréquences avaient de meilleures capacités à traiter l'information basse fréquence par rapport à des sujets qui ne présentaient pas de ZIC. Selon Vestergaard (2004 ; p125), ces patients auraient même développé des capacités supra-normales pour le traitement des indices basse-fréquence de la parole. Des études beaucoup plus anciennes avaient déjà fait des constats similaires (par exemple Rhodes en 1966), mais à notre connaissance aucune étude n'a essayé de relier les performances de ces patients à des facteurs autres que l'audibilité ou que la présence/absence de ZIC.

Dans notre étude, le TEN test (test de détection des ZIC) n'a pas été utilisé puisque selon Moore (2001) qui a mis le test en place, lorsque la perte est supérieure à 80-90 dB HL nous suspectons une ZIC, or tous nos patients avaient au moins ce degré de perte dans les hautes fréquences et avaient probablement des ZIC. Moore et al. (2004) ajoutent que pour de tels degrés de surdité, le TEN test n'est pas réalisable car l'intensité à laquelle il faudrait mettre le bruit masquant n'est pas toléré par les sujets.

Notre étude a été réalisée chez 20 sujets malentendants présentant une surdité supérieure à 80 dB au-delà de 1000 Hz avec des graves relativement bien conservés. Ils avaient des durées de surdité allant de 4 à 64 ans. Nous les avons classés en 3 catégories de surdité : surdité précoce (congénitale ou apparue avant l'âge de 4 ans), surdité évolutive et surdité brusque.

Le résultat principal de cette étude montre une corrélation négative significative entre l'âge d'apparition de la surdité et le score de reconnaissance des syllabes, ainsi qu'une corrélation positive entre la durée de surdité et les performances. Nous avons également trouvé, comme il était prévisible, une corrélation significative entre les performances et l'audibilité. En moyenne, nous trouvons une amélioration des performances de 1,4% par année de surdité, ce qui signifie qu'avec un audiogramme équivalent, un sujet qui est sourd depuis 30 ans a un score de compréhension de 28% supérieur au score d'un sujet qui serait sourd depuis 10 ans. Par ailleurs, nous avons mis en évidence que les surdités congénitales ou de la petite enfance comprenaient significativement mieux que les sujets qui avaient une surdité acquise (évolutive ou brusque).

Nos résultats indiquent que chez ces malentendants il existe des facteurs, au-delà de l'audibilité, qui peuvent sévèrement affecter l'intelligibilité de la parole.

Relationship Between Age of Hearing-Loss Onset, Hearing-Loss Duration, and Speech Recognition in Individuals with Severe-to-Profound High-Frequency Hearing Loss

FABIEN SELDRAN^{1,2}, STÉPHANE GALLEG^{1,4}, CHRISTOPHE MICHEYL³, EVELYNE VEUILLET¹, ERIC TRUY^{1,4}, AND HUNG THAI-VAN^{1,4}

¹CNRS UMR 5020, Laboratoire 'Neurosciences Sensorielles Comportement Cognition', Université Claude Bernard Lyon I, 50 Avenue Tony Garnier, Lyon Cedex 07 69366, France

²Vibrant Med-El Hearing Technology GmbH, 400 Ave. Roumanille, BP 309, Sophia-Antipolis 06906, France

³Auditory Perception and Cognition Laboratory, Department of Psychology, University of Minnesota, Elliott Hall, 75 East River Road, Minneapolis, MN 55455, USA

⁴Audiology and ENT Department, Edouard Herriot Hospital, 3 Place d'Arsonval, Lyon 69437, France

Received: 10 October 2010; Accepted: 7 February 2011; Online publication: 25 February 2011

ABSTRACT

The factors responsible for interindividual differences in speech-understanding ability among hearing-impaired listeners are not well understood. Although audibility has been found to account for some of this variability, other factors may play a role. This study sought to examine whether part of the large interindividual variability of speech-recognition performance in individuals with severe-to-profound high-frequency hearing loss could be accounted for by differences in hearing-loss onset type (early, progressive, or sudden), age at hearing-loss onset, or hearing-loss duration. Other potential factors including age, hearing thresholds, speech-presentation levels, and speech audibility were controlled. Percent-correct (PC) scores for syllables in disyllabic words, which were either unprocessed or lowpass filtered at cutoff frequencies ranging from 250 to 2,000 Hz, were measured in 20 subjects (40 ears) with severe-to-profound hearing losses above 1 kHz. For comparison purposes, 20 normal-hearing subjects (20

ears) were also tested using the same filtering conditions and a range of speech levels (10–80 dB SPL). Significantly higher asymptotic PCs were observed in the early (≤ 4 years) hearing-loss onset group than in both the progressive- and sudden-onset groups, even though the three groups did not differ significantly with respect to age, hearing thresholds, or speech audibility. In addition, significant negative correlations between PC and hearing-loss onset age, and positive correlations between PC and hearing-loss duration were observed. These variables accounted for a greater proportion of the variance in speech-intelligibility scores than, and were not significantly correlated with, speech audibility, as quantified using a variant of the articulation index. Although the lack of statistical independence between hearing-loss onset type, hearing-loss onset age, hearing-loss duration, and age complicate and limit the interpretation of the results, these findings indicate that other variables than audibility can influence speech intelligibility in listeners with severe-to-profound high-frequency hearing loss.

Correspondence to: Fabien Seldran · CNRS UMR 5020, Laboratoire 'Neurosciences Sensorielles Comportement Cognition' · Université Claude Bernard Lyon I · 50 Avenue Tony Garnier, Lyon Cedex 07 69366, France. Telephone: +33-4-72110503; fax: +33-4-72110504; email: fseldran@yahoo.fr

Keywords: high-frequency hearing loss, hearing-loss onset age, hearing-loss duration, speech intelligibility

INTRODUCTION

Hearing-impaired (HI) individuals often vary considerably in their ability to understand speech. The factors responsible for this interindividual variability are not entirely clear. While audibility plays an important role (Dubno et al. 1989b; Rankovic 1991; Hogan and Turner 1998; Rankovic 1998, 2002), listeners with severe hearing losses have often been found not to benefit from amplification, even when audibility calculations indicate that they should (Pavlovic 1984; Kamm et al. 1985; Dubno et al. 1989a; Ching et al. 1998; Hogan and Turner 1998; Baer et al. 2002).

Studies of cochlear-implant (CI) subjects indicate that, for these subjects, whether hearing-loss occurred before or after the acquisition of language, and how long the duration of deafness was prior to implantation, are two of the most significant predictors of speech-recognition performance (Tong et al. 1988; Busby et al. 1992, 1993; Dawson et al. 1992; Gantz et al. 1993; Hinderink et al. 1995; Kessler et al. 1995; Shipp and Nedzelski 1995; Blamey et al. 1996; Okazawa et al. 1996; Rubinstein et al. 1999; Van Dijk et al. 1999; Friedland et al. 2003; Leung et al. 2005; Green et al. 2007; Gantz et al. 2009). In general, the results show that individuals in whom hearing-loss occurred postlingually achieve statistically higher speech-recognition scores than individuals who became deaf prior to the acquisition of language (Tong et al. 1988; Busby et al. 1993; Hinderink et al. 1995; Okazawa et al. 1996), and that the duration of deafness correlates negatively with speech-recognition performance (Gantz et al. 1993; Blamey et al. 1996; Rubinstein et al. 1999; Gantz et al. 2009). These findings have been variously explained in terms of neural plasticity and detrimental effects of sensory deprivation versus beneficial effects of sensory stimulation on auditory pathways involved in speech recognition, and in terms of perceptual learning of speech cues.

While these findings concern specifically CI subjects, age at hearing-loss onset and hearing-loss duration may influence speech-recognition performance in non-implanted HI individuals as well. Firstly, hearing losses incurred before the acquisition of language may impede the development of speech-recognition skills, so that subjects with congenital or early-onset hearing loss are less likely to achieve as high levels of speech understanding as individuals with similar hearing thresholds but later-onset hearing loss. Conversely, individuals in whom the hearing loss is already in place before language acquisition may develop long-term memory traces (or “templates”) that are more adapted to decoding the impoverished and distorted speech signals that they receive. As a result, these individuals might be better able to recognize speech than subjects who have similar audiograms, but who incurred a hearing loss at a later age, at which point their central auditory system

was perhaps less plastic. In addition, speech-recognition performance could be related statistically to the age or type (early versus late) of hearing-loss even if it does not directly depend on these variables. For instance, early hearing losses might be more frequently associated with certain types of peripheral or central damage, which impact speech-perception performance in a different way than other types of peripheral damage more frequently associated with late-onset losses. Similarly, hearing-loss duration could be a factor of speech-recognition performance if, for example, having a relatively stable or slow-changing hearing-loss allowed HI individuals to progressively adjust their listening strategies to optimally decode the impoverished and distorted speech signals that they receive (Rhodes 1966; Niemeyer 1972). On the other hand, progressive degeneration of neural pathways following peripheral insults could result in poorer speech-recognition performance in the individuals in whom a hearing-loss has been present for several years than in individuals in whom peripheral damage was incurred more recently, even if these individuals have similar pure-tone audiograms (Kujawa and Liberman 2009).

Accordingly, this study sought to test whether hearing-loss onset age, onset type, and hearing-loss duration can contribute to explain (in a statistical sense) the large variability in speech-recognition scores in individuals with severe-to-profound high-frequency hearing loss.

METHODS

Subjects

Twenty HI subjects (11 female, nine male; ages, 17–71 years; median age=44.5 years) and 20 normal-hearing (NH) subjects (seven female, 13 male; ages, 20–32 years; median age=23 years) took part in the study. All NH subjects had pure-tone hearing thresholds lower than 20 dB HL at octave frequencies between 250 and 8,000 Hz. The HI subjects had severe-to-profound hearing loss at and above 1 kHz. Most of them also had some degree of hearing loss (ranging from mild to profound) below 1 kHz. Figure 1 shows the mean and individual hearing thresholds of the HI listeners between 250 and 4,000 Hz, separately for the left and right ears. For the purpose of this study, thresholds were measured in half-octave steps between 250 and 2,000 Hz; these frequencies were the same as the lowpass-filter cutoff frequencies (CFs) used in the speech tests described below. The downward-pointing triangles in Figure 1 indicate “unmeasurable” thresholds. Thresholds were considered unmeasurable when the subject did not detect the pure-tone stimulus at the maximum sound level that could be produced by the audiometer at the considered test frequency, with contralateral masking to avoid cross-over effects. This maximum level varied from

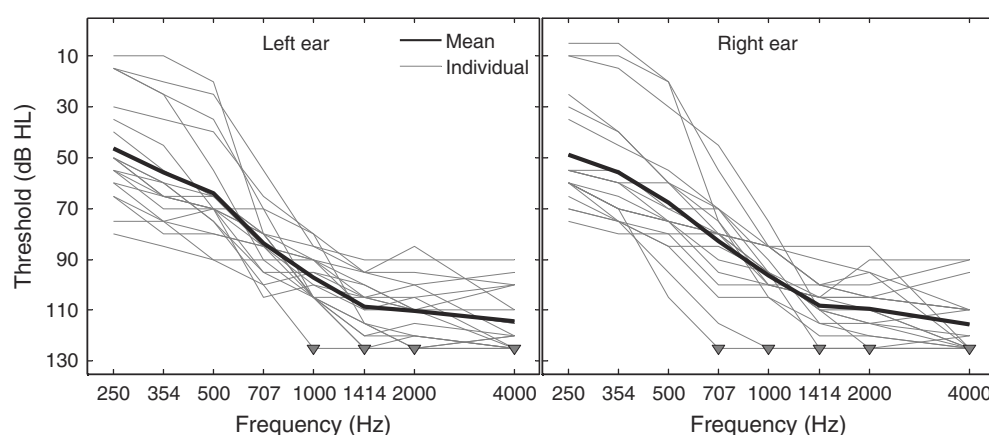


FIG. 1. Mean and individual pure-tone hearing thresholds in HI subjects. Thick lines, mean thresholds across all HI subjects. Light lines, individual data. Downward-pointing triangles indicate unmeasurable thresholds.

105 to 120 dB HL depending on the test frequency and channel. It was equal to 120 dB HL between 600 and 2,000 Hz. For data-visualization and data-analysis purposes, unmeasurable thresholds were set arbitrarily to a value of 125 dB HL. Three of the 40 ears tested in this study had unmeasurable thresholds at 1 kHz. At 4 kHz, 17 ears had unmeasurable thresholds. Air-bone gaps were never greater than 10 dB at any of the frequencies tested, indicating that the hearing losses were sensorineural rather than conductive.

Table 1 provides additional individual information, including gender, age, and pure-tone threshold averages over two frequency ranges (250–710 Hz and 1–4 kHz). Based on the information provided by the subjects concerning the history of their hearing loss and, when available, audiometric records, ears were categorized into three groups corresponding to three types of hearing-loss onset: “early” (ten ears), which was defined as congenital hearing loss or hearing loss incurred at or before the age of 4, with little or no change in pure-tone thresholds thereafter; “progressive” (23 ears), which was defined as hearing loss having developed over the course of several years (in some cases, several decades), after the age of 4; and “sudden” hearing loss (seven ears), which was defined as abrupt hearing loss having occurred after the age of 4 (to avoid overlap with the “early” group). With one exception (subject 18), the type of hearing loss was the same for both ears in every subject; that subject exhibited a sudden and relatively stable hearing loss in the right ear, but a more progressive hearing loss in the left ear. Subjects 13 and 15 were placed into the “progressive” category even though hearing loss was diagnosed within their first year of life, because their hearing loss progressed over the course of several years, well beyond the fourth year of life. We acknowledge that these grouping criteria are somewhat arbitrary, and that any conclusion concerning the effects of hearing-loss type based on this grouping must be qualified accordingly. Importantly, the subjects in these three groups were selected in such a way that

they had similar age characteristics (“early”: 17–64 years, median=42 years; “progressive”: 30–71, median=47 years; “sudden”: 37–63, median=37), and similar pure-tone thresholds within the 250–4,000 Hz range (Fig. 2). There was no significant difference in pure-tone thresholds or age between the groups (group \times ear \times frequency analysis of variance (ANOVA) on thresholds with group as a between-subject factor and ear and frequency as within-subject factors: $F_{2, 17}=0.386$ and $P=0.686$; one-way ANOVA on age with group as a between-subject factor: $F_{2, 19}=0.73$ and $P=0.496$).

Since the three groups of hearing-impaired listeners did not differ significantly with respect to age but did differ with respect to hearing-loss onset age (which was lower, by definition, for the early-onset group than for the other two groups), we expected them to also differ with respect to hearing-loss duration, since this variable was equal to the listeners’ age minus the hearing-loss onset age. Due perhaps to interindividual variability, the main effect of hearing-loss duration in a two-way ANOVA with group as a between-subject factor and ear as a within-subject factor failed to reach statistical significance ($F_{2, 17}=3.19$ and $P=0.067$). However, a planned contrast analysis comparing hearing-loss duration in the early-onset group versus the mean hearing-loss duration across the other two groups showed a significant difference ($t=18.21$ and $P=0.022$). Hearing-loss duration was longer, by 18 years on average, in the early-onset group than in the other two groups. Accordingly, tests were performed to determine whether differences in hearing-loss duration could account for differences in speech-recognition scores across the three groups; these tests are described in “Data analysis” and “Results”.

Table 1 additionally lists the duration of the hearing loss, which was determined by subtracting the age at which hearing-loss was diagnosed or suspected to have started, as indicated in the subject’s history, from the subject’s age at the time of testing. For 14 out of the 20 subjects, the cause of the hearing loss was unknown. For the remaining six subjects, the

TABLE 1									
Individual participant information									
Gender	Age (years)	Ear	Hearing loss		Duration (years)	HA use (years)	Speech level (dB SPL)	PTA 250–710 Hz (dB HL)	PTA 1–4 kHz (dB HL)
			Onset type	Duration (years)					
1	M	R	Early	17	14	100	64	98	
		L	Early	17	14	100	65	93	
2	M	R	Early	28	18	100	79	104	
		L	Early	28	18	100	76	101	
3	M	R	Early	39	38	100	75	98	
		L	Early	39	38	100	71	105	
4	F	R	Early	50	45	100	70	110	
		L	Early	50	45	100	68	109	
5	F	R	Early	64	10	80	25	108	
		L	Early	64	0	80	36	119	
6	F	R	Progressive	4	4	100	63	118	
		L	Progressive	4	4	90	73	120	
7	M	R	Progressive	9	3	100	53	96	
		L	Progressive	9	3	100	64	100	
8	F	R	Progressive	15	0	90	65	110	
		L	Progressive	15	5	90	65	118	
9	F	R	Progressive	15	0	80	24	96	
		L	Progressive	15	0	90	29	89	
10	M	R	Progressive	18	14	100	71	95	
		L	Progressive	18	14	100	78	95	
11	F	R	Progressive	18	17	100	84	110	
		L	Progressive	18	5	100	89	108	
12	F	R	Progressive	28	18	100	76	100	
		L	Progressive	28	18	100	83	100	
13	M	R	Progressive	31	21	80	26	105	
		L	Progressive	31	21	80	30	104	
14	F	R	Progressive	35	2	100	90	125	
		L	Progressive	35	2	90	46	118	
15	F	R	Progressive	38	11	100	83	116	
		L	Progressive	38	11	100	65	111	
16	M	R	Progressive	41	18	100	51	111	
		L	Progressive	41	18	100	60	116	
17	F	R	Sudden	6	0	80	49	108	
		L	Sudden	6	0	100	71	124	
18	M	R	Sudden	22	12	100	74	118	
		L	Progressive	22	12	95	71	115	
19	F	R	Sudden	26	9	100	68	99	
		L	Sudden	26	0	100	69	100	
20	M	R	Sudden	7	7	100	86	125	
		L	Sudden	44	3	100	43	110	

The numbers in the first column refer to subject numbers. “HA use” refers to the number of years of hearing-aid use. “PTA 250–710 Hz” and “PTA 1–4 kHz” indicate pure-tone thresholds averaged over the indicated frequency range

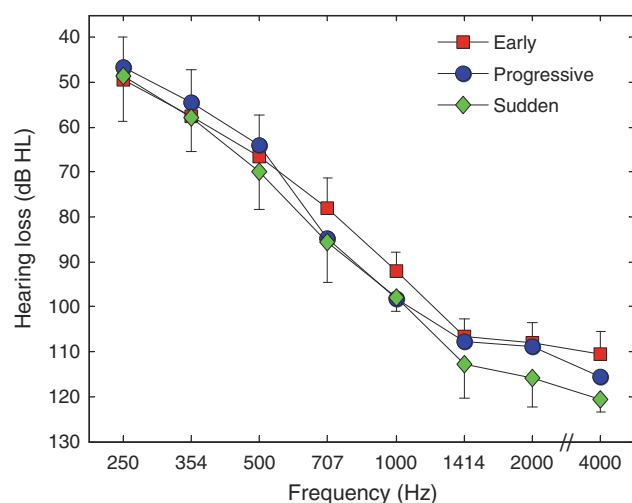


FIG. 2. Mean pure-tone hearing thresholds for the three groups of hearing-loss onset type. Color is used to indicate group, as indicated in the key. The same group color coding scheme is used in all figures: *red* for the early-onset group; *blue* for the progressive-onset group; *green* for the sudden-onset group. Error bars show standard error of the mean across ears within each group. To avoid clutter, error bars are only shown for the lowest and highest mean at each CF.

suspected causes of hearing loss were as follows: malformation of the vestibular aqueduct (subject 1), electrocution (subject 3), Pott's disease (subject 5), Usher's syndrome (subject 15), otitis media or genetic (subject 19), and head injury (subject 20).

At the time of testing, all but six of the subjects were wearing hearing aids in both ears on a regular basis in their daily life. Out of the six subjects who were not wearing hearing aids in both ears on a regular basis, one (subject 9) was never fitted, one (subject 13) had been fitted bilaterally but did not wear his hearing aids on a regular basis, one (subject 17) tried using hearing aids twice but never succeeded in wearing them for a long period of time, and the remaining three were wearing a hearing aid in one ear only (the right ear for subjects 5 and 19, the left ear for subject 8). The number of years of hearing-aid use for each ear is indicated in Table 1. The subjects did not wear their hearing aids during any of the tests performed in this study.

In accordance with the Declaration of Helsinki, written informed consent was obtained from the subjects prior to their inclusion into the study. The study was approved by the local Ethics Committee (CPP Sud-Est IV, Centre Leon Berard de Lyon, France, no. ID RCB: 2008-A01479-46).

Stimuli and procedure

Speech intelligibility was measured using 40 lists of ten disyllabic French words (Fournier 1951). These lists may be regarded as the French-language equivalent of the American Spondaic lists. The words were uttered by a male talker and recorded on a CD, sampled at 44.1 kHz,

with a 16-bit quantization range. They were lowpass filtered digitally in the frequency domain using the FFT filter of the Adobe Audition software, with a stopband attenuation of 70 dB. Seven lowpass CFs, corresponding to half-octave steps between 250 and 2,000 Hz (i.e., 250, 354, 500, 707, 1,000, 1,414, and 2,000 Hz) were used, yielding a total of 8 test conditions (the seven lowpass-filtering conditions plus the "unprocessed" condition), for a total of 3,200 stimuli (corresponding to the eight processing conditions \times 40 lists \times ten words per list). To facilitate the description of the results, hereafter and in the figures, the "unprocessed" condition is referred to as another CF condition, with CF=22,050 Hz, i.e., half of the rate at which the digitized word lists were sampled.

The rationale for testing multiple CF conditions was as follows. Firstly, we reasoned that relationships between speech-recognition performance and other variables of interest in this study, such as hearing-loss duration or hearing-loss onset type, might be CF-dependent. Previous studies in which intelligibility has been measured for lowpass-filtered speech in listeners with high-frequency hearing losses (e.g., Murray and Byrne 1986; Hogan and Turner 1998; Vickers et al. 2001) have found that in some listeners, performance varied nonmonotonically with CF. In these listeners, surprisingly, the wideband condition did not always yield the highest performance. While the reasons for this effect are not completely known, one hypothesis is that, in frequency regions where the hearing loss is severe or profound, the speech signal received by the central auditory system is so distorted that it provides no useful information to the listener, and actually interferes with the processing of speech information contained in other (usually, lower) frequency regions, where the hearing loss is less marked. The fact that not all listeners show evidence of this effect may be due, not just to interindividual differences in hearing-loss characteristics (as reflected in pure-tone thresholds and audibility calculations), but also, to more central factors, such as differences in the ability to "ignore" peripheral channels in which information is too distorted to be useful, or a better ability to extract information from a distorted peripheral signal. To the extent that these factors are related to hearing-loss onset age or to hearing-loss duration, correlations between these variables and speech-recognition performance should be observed specifically in wideband conditions, or in conditions where the CF of the lowpass filter is sufficiently high for the speech signal to contain energy in frequency regions where peripheral information is highly distorted. In addition, testing multiple filtering conditions yielded a wider range of performance levels within a given listener, and increased the likelihood of capturing relationships between performance and other variables. A third reason for testing multiple lowpass-filtering conditions in HI listeners was that several of these listeners were candidates for implantation with a

short-electrode implant and electro-acoustic stimulation. Therefore, it seemed both interesting and potentially important to measure the ability to understand speech for different hypothetical scenarios of preservation of low-frequency acoustic hearing in these listeners (Moore et al. 2010), and to investigate whether this ability was related to hearing-loss duration or hearing-loss onset type. For NH listeners, testing multiple CF conditions allowed us to derive frequency-importance functions specific to the speech stimuli that were used in this study. These frequency-importance functions were then used to compute articulation-index (AI) predictions in HI listeners, as explained in the next section.

For HI listeners, the “unprocessed” (CF=22,050 Hz) condition was tested first, at three or four different sound levels: 70, 80, 90, and/or 100 dB SPL, where SPL refers to the root-mean-square (RMS) sound-pressure level of the stimuli across all stimuli. The highest level (100 dB SPL) was only tested in listeners who did not find it uncomfortably loud. In this initial test, and in all subsequent tests, the left and right ears of each subject were tested, separately and in random order. During this initial phase of the test, the subjects were asked to indicate which listening level they found most comfortable. These levels are listed in Table 1. One of these subjects described the intensity of 90 dB SPL as too low, and that of 100 dB SPL as too loud; for this listener, an intermediate level (95 dB SPL) was used. On average, the level chosen by the subjects as most comfortable was approximately equal to 95 dB SPL. Importantly, average speech-presentation levels did not differ significantly between the three hearing-loss onset groups ($F_{2, 17} = 0.108$ and $P = 0.898$). The average speech-presentation levels were equal to 96 dB (SD=8.44) for the “early” group, 95 dB (± 7.30) for the “progressive” group, 97 dB (± 6.67) in the “sudden” group.

The comfortable level chosen by the subject was kept during subsequent testing, which involved low-pass-filtered speech. However, because the bandwidth of the filtered stimuli decreased with CF, and was always smaller than for the “unprocessed” stimuli, the actual SPL of the filtered stimuli was lower than the SPL of the unprocessed (CF=22,050 Hz) stimuli, and it decreased with CF—as was the case in previous studies (e.g., Hogan and Turner 1998; Vickers et al. 2001). The NH listeners were tested using eight speech levels, ranging from 10 to 80 dB SPL in steps of 10 dB, for each of the CF conditions.

For HI listeners, two lists of ten words were drawn at random for each stimulus condition. The lists were presented monaurally to the listener, whose task was to repeat the words as they heard them. Because the number of conditions tested in NH listeners was larger (due to the testing of eight speech levels for each CF), in these listeners, only one randomly selected list of ten words was presented in each stimulus condition, and

only the right ear was tested. The different lowpass-filtering conditions were tested in the same order in all subjects, starting with the highest CF condition (22,050 Hz) first, then proceeding with the next highest CF (2,000 Hz), and so forth, until the lowest CF (250 Hz) was tested. Our decision to adopt this reverse-order testing procedure, instead of a completely randomized testing procedure, was based on the following considerations. Firstly, we reasoned that presenting all stimuli corresponding to a given filtering condition as a group, rather than interspersed randomly amidst stimuli filtered differently, would alleviate the detrimental influence of lack of familiarity with the filtered stimuli. Secondly, we reasoned that testing the conditions in order of increasing difficulty would promote higher performance in the most difficult conditions, by allowing listeners to adapt progressively to increasing task difficulty. In an attempt to avoid fatigue and attentional lapses, the subjects were encouraged to take breaks whenever they felt the need to do so during the test.

Data analysis

Speech-recognition performance was measured by counting the number of syllables that were correctly repeated by the listener in a given test condition, dividing this number by the total number of syllables presented in that condition (40 for HI listeners and 20 for NH listeners), and multiplying the result by 100 to obtain a percent-correct (PC) score. Our decision to count correctly repeated syllables, rather than correctly repeated words, was motivated by the observation that listeners could often identify correctly one of the two syllables in the word, but were not able to repeat the whole word. In addition, we reasoned that scoring syllables would result in more accurate estimates of PC than scoring whole-words since there were more syllables than words, and the joint probability of correctly identifying both syllables in a word was not equal to the product of the corresponding marginal probabilities.

Statistical relationships between PC and other variables, including hearing-loss duration, hearing-loss onset age, age, and pure-tone thresholds, were investigated using Pearson’s correlation coefficients. To take into account the lack of statistical independence between variables corresponding to the left and right ears in the same subject, the sampling distribution of Fisher’s z -transformed correlation coefficients (Fisher 1915) under the null hypothesis (no correlation) was computed using a statistical resampling technique (bootstrap) (Efron and Tibshirani 1994). This involved computing the test statistic repeatedly, a large number of times, with the data shuffled randomly on each trial. Measures from the left and right ears were kept in pairs at all times, so that correlations between the two ears were taken into account. Between-ear correlations were

reflected in a wider distribution of the test statistic, and in larger confidence intervals, compared with the uncorrelated case. Partial-correlation coefficients (Fisher 1924) were used to test statistical relationships between two variables (e.g., hearing-loss duration and PC), taking into account correlations between these variables and a third variable (e.g., age).

To determine whether hearing-loss type was a significant factor of performance, we performed ANOVA on the PC values measured in HI listeners, using hearing-loss type (early, progressive, and sudden) as an across-subject factor, and CF and ear (left–right) as within-subject factors. To determine whether hearing-loss duration was a factor in differences in PC between the groups, it was entered as a covariate in the ANOVA—a procedure commonly known as “analysis of covariance” (ANCOVA). Because hearing-loss onset type could not be used as both an across-subject factor and a within-subject (across-ear) factor, for the purpose of these two analyses (and these two analyses only), subject 18, who had different types of hearing-loss onset for the left and right ears, was placed into the “sudden” group. However, we verified that the main conclusions of these two analyses were qualitatively unchanged when that subject was placed into the “progressive” group instead. These ANOVA and ANCOVA were preceded by a Mauchly’s test of sphericity, and a Huynh–Feldt correction was applied when required. Multiple post hoc comparisons following the finding of significant main effects in the ANOVA were performed according to Fisher’s protected least-significant-difference (LSD) procedure. All statistical tests were performed using untransformed PC data. However, we verified that the conclusions remained unchanged when the analyses were performed instead on arcsine-transformed data.

The influence of speech audibility on the results was evaluated by computing an index inspired from the AI (ANSI 1969), which we denote AI'. To this aim, we used the PC data obtained for different CFs and levels in the NH listeners to, firstly, derive frequency-importance functions for the speech stimuli used in this study. Frequency-importance functions represent the relative importance of different frequency bands for speech understanding (see, e.g., French and Steinberg 1947). Deriving importance functions for the specific speech stimuli used in this study was necessary because there was no guarantee a priori that the “standard” importance functions for English, which can be found in, e.g., ANSI (1997), were the same as the importance functions for the French dissyllabic word lists used here. However, we also performed similar calculations using the standard importance functions provided in ANSI (1997) and found that, although predicted AI values based on these standard functions were generally lower than those obtained using the specific importance

functions measured in this study, our main conclusions regarding the relationship between speech audibility and other variables did not depend critically on which of these two functions was used.

To derive frequency-importance functions for the word lists used in this study, we first transformed the PCs measured in NH listeners into d' . This was done by solving the following equation for d' numerically.

$$PC = 100 \int_{-\infty}^{+\infty} \phi(z - d') \Phi^{m-1}(z) dz. \quad (1)$$

In this equation, which follows from the equal-variance Gaussian signal-detection-theory model for the 1-of- m identification task (Green and Dai 1991; Macmillan and Creelman 2005), ϕ and Φ denote the standard normal and cumulative standard normal functions, respectively, and m , the number of response alternatives, was set to 288, the number of possible pairwise combinations of 16 vowels and 18 consonants. The rationale for applying this nonlinear transformation is that, unlike PC, d' provides an additive measure of band importance. The average importance of frequency band i , $1 < i \leq n$, was estimated as

$$I_i = \frac{d'_i - d'_{i-1}}{\sum_{j=1}^n I_j}. \quad (2)$$

where d'_{i-1} and d'_i denote the mean d' s measured in the $(i-1)$ th and i th CF conditions (e.g., 707 Hz and 1,000 Hz), respectively, across all NH listeners and speech levels, and n is the total number of frequency bands (which was equal to the number of CF conditions). For the lowest frequency band (band 1),

$$I_1 = \frac{d'_1}{\sum_{j=1}^n I_j}. \quad (3)$$

The denominator in Eqs. 2 and 3 ensures that the total importance across all n frequency bands equals 1 (French and Steinberg 1947).

To estimate the audibility of the speech signals presented to the HI listeners, we used a simplified version of the formulas described in ANSI (1997), which was similar to the formula originally proposed in ANSI (1969). Specifically, the mean audibility of the speech signals for band i was computed as:

$$A_i = \max(\min(L_i + L_0 - T_i - N_i + C, 30), 0)/30 \quad (4)$$

In this equation, L_i (dB SPL/Hz) denotes the long-term average spectrum (LTAS) level of the speech in band i , assuming a 0-dB overall SPL; L_0 (dB SPL) denotes the actual overall SPL of the speech signals; T_i denotes the absolute threshold (dB HL) at the center

frequency of band i , which was estimated by interpolating the pure-tone hearing thresholds measured in the HI listeners; N_i denotes the “reference internal noise spectrum level” for band i , as defined in the ANSI (1997) standard. The max and min operations, and the division by 30, ensure that the result is comprised between 0 and 1. Finally, the AI-like measure, AI' , was computed as,

$$AI' = \sum_{i=1}^n I_i A_i \quad (5)$$

The value of the constant, C , in Eq. 4 was originally set to 15, as in ANSI (1997). However, a value of 25 was subsequently found to yield a better overall fit between the psychometric functions for the NH subjects, which were used to determine the specific frequency-band importance functions used in the current study, and the corresponding psychometric functions, which were predicted based on AI' using the approach described below; therefore, the value of 25 was kept.

To derive PC predictions based on the computed AI' values, we performed linear regression on the AI' -versus- d' data of the NH subjects. We then used the resulting linear-regression equation to transform the AI' values obtained in the HI subjects into d' . Finally, the resulting d' values were transformed into PC using Eq. 1.

Apparatus

Pure-tone audiometry was performed using a MADSEN Orbiter 922 audiometer and TDH39 earphones. For the speech tests, the signals were burned on CDs and played via a CD player (PHILIPS CD723) connected to the MADSEN audiometer. The signals were presented monaurally to the listeners via TDH39 earphones.

RESULTS

Speech intelligibility as a function of CF and effect of hearing-loss onset type

Figure 3 shows mean PC as a function of CF for NH listeners. As expected, PC increased with CF. It also increased with level up to 40 dB SPL, beyond which no consistent improvement in PC with level was observed. Accordingly, PC values for each CF were averaged across the highest five speech levels (40–80 dB SPL). The lines through the data show best-fitting psychometric functions computed as described in the Appendix. These psychometric functions are re-plotted in the next figure to facilitate comparisons between NH and HI listeners.

Figure 4 shows mean PC as a function of CF for HI listeners. The results of a group \times ear \times CF ANOVA with group as a between-subject factor and ear and CF as

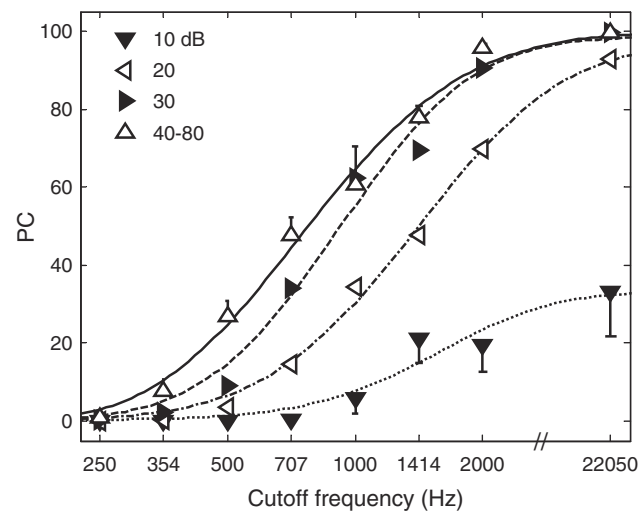


FIG. 3. PC as a function of lowpass-filter CF in NH subjects. The different combinations of symbols and line styles indicate different speech levels, as indicated in the key. The curves corresponding to levels comprised between 40 and 80 dB SPL were averaged because they were similar and largely overlapped. Error bars show standard error of the mean across subjects. To avoid clutter, error bars are only shown for the lowest and highest mean at each CF condition.

within-subject factors showed a significant main effect of group ($F_{2, 17}=4.61$ and $P=0.025$), and a significant interaction between hearing-loss type and CF (Huynh–Feldt-corrected $F_{14, 119}=2.10$, and $P=0.036$). Post hoc

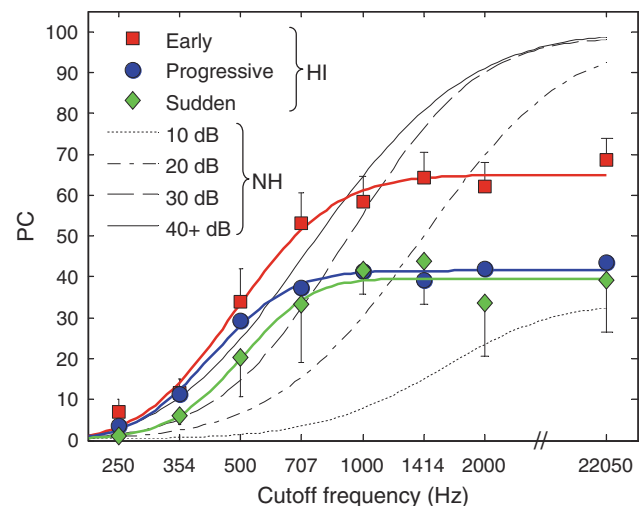


FIG. 4. PC as a function of lowpass-filter CF in HI subjects. The data corresponding to the early-, progressive-, and sudden-onset groups are indicated by different colors, using the same scheme as in Figure 2. To facilitate comparisons, the best-fitting psychometric functions of the NH subjects are re-plotted (from Fig. 3) as black lines, with the different line styles corresponding to different speech-presentation levels (as in Fig. 3), as indicated in the key. The psychometric functions for speech-presentation levels of 40 dB SPL or higher were largely overlapped, and were averaged before display, to avoid clutter; the corresponding, averaged psychometric function is shown, labeled as 40+ dB. Error bars show standard error of the mean across ears within each group. To avoid clutter, error bars are only shown for the lowest and highest mean at each CF.

tests (Fisher's LSD) comparing the three groups two at a time showed significant differences in PC between the "early" and "progressive" groups (mean difference = 12.94 and $P=0.032$), as well as between the "early" and "sudden" groups (mean difference = 20.01 and $P=0.010$), but not between the "progressive" and "sudden" groups (mean difference = 7.07 and $P=0.254$). Contrast analyses, which compared PC in the "early" group versus mean PC across the other two groups for each CF separately showed significant differences for CFs of 1,000 Hz and higher (1,000 Hz: mean difference = 18.61 and $P=0.038$; 1,414 Hz: mean difference = 23.98 and $P=0.020$; 2,000 Hz: mean difference = 25.21 and $P=0.005$; 22,050 Hz: mean difference = 28.43 and $P=0.005$), but not for lower CFs (250 Hz: mean difference = 4.68 and $P=0.072$; 354 Hz: mean difference = 2.75 and $P=0.457$; 500 Hz: mean difference = 9.45 and $P=0.243$; 707 Hz: mean difference = 18.68 and $P=0.065$).

As mentioned above, the three HI groups did not differ significantly with respect to age, pure-tone thresholds, or speech-presentation levels. However, as mentioned above, duration of hearing-loss was significantly longer on average for the early-onset group than for the other two groups. To determine whether this difference in hearing-loss duration could account for the differences in PC between the groups, hearing-loss duration was entered as a covariate in a group \times ear \times CF ANOVA, with group as a between-subject factor and ear and CF as within-subject factors. This analysis (ANCOVA) showed no significant difference between the groups ($F_{2, 16} = 1.84$ and $P=0.192$), and no significant effect of the cofactor ($F_{1, 16} = 3.03$ and $P=0.101$). This outcome indicates that the differences in PC between the HI groups could be accounted for by differences in hearing-loss duration. The three HI groups also differed with respect to mean age at hearing-loss onset, which was lower for the "early" group (1.6 years) than for the "progressive" and "sudden" groups (25.82 and 28.14 years, respectively). However, since hearing-loss onset age was one of the criteria used to define the groups, this variable could not be meaningfully used as a cofactor in the ANOVA.

Lastly, it is interesting to consider whether speech audibility can account for differences between the three HI groups, as well as between HI and NH listeners. The upper panel in Figure 5 shows mean AI' as a function of CF for the three HI groups and for the NH listeners. The lower panel shows the mean PC predicted based on AI' , as a function of CF, for the three HI groups and for the NH listeners. As can be seen, the mean AI' and predicted PC functions of the three groups of HI listeners are generally close to each other. The 15- to 25-percentage-points difference between the "early" group and the two other HI groups that was observed for CFs of 1 kHz and higher in Figure 4, is not reflected in the predicted-PC functions (Fig. 5, lower panel). This

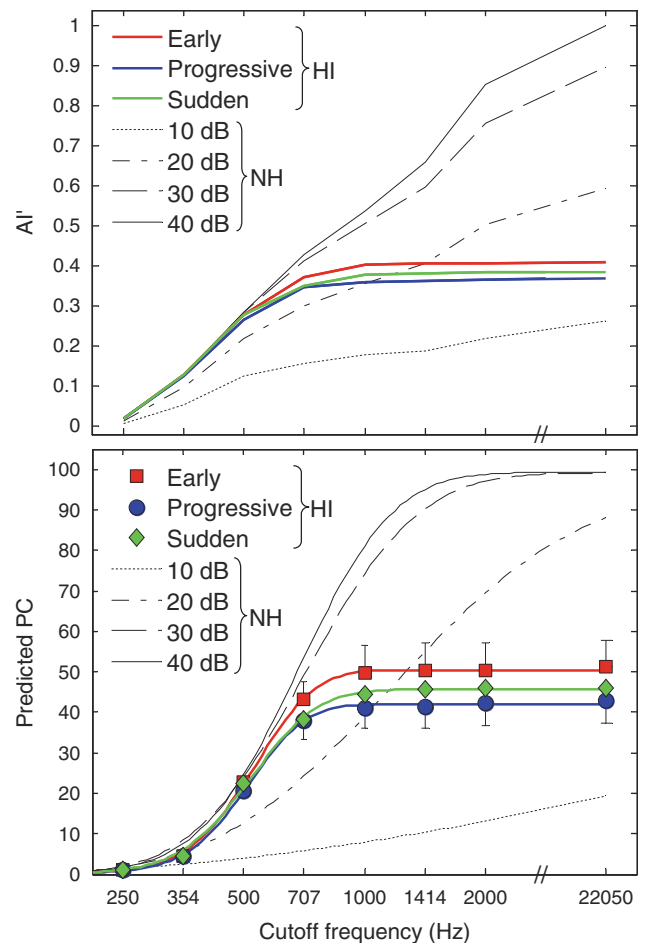


FIG. 5. AI' and predicted PC as a function of CF for HI and NH subjects. Upper panel, AI' . Lower panel, PC predicted based on AI' . For HI subjects, AI' (for the upper panel) or predicted PC values (for the lower panel) were averaged across ears within each hearing-loss onset group. As in Figures 2 and 4, the different hearing-loss onset groups are indicated by different colors. For NH subjects, AI' (for the upper panel) or predicted PC values (for the lower panel) were averaged across subjects separately for each speech-presentation level. The AI' and PC functions for speech-presentation levels higher than 40 dB SPL were essentially identical to the psychometric function corresponding to 40 dB SPL. The data of the NH subjects are indicated by black lines, with different line styles indicating different speech-presentation levels. AI' curves for levels higher than 40 dB SPL were identical to the curve corresponding to 40 dB SPL.

suggests that the superior speech-recognition performance of the "early" group at high CFs cannot be explained simply in terms of audibility; rather, it appears that, for these CFs, the listeners of the "early" group were better able to extract useful information from the speech signals than the listeners in the other two HI groups, on average.

Correlations between speech-recognition performance and other variables

Statistically significant correlations between PC and hearing-loss onset age, and between PC and hearing

TABLE 2

Correlation coefficients between PC and hearing-loss onset age, hearing-loss duration, or age, and corresponding *P* values

CF	Onset age		Duration		Age	
	<i>r</i>	<i>P</i>	<i>r</i>	<i>P</i>	<i>r</i>	<i>P</i>
250	−0.37	0.035	0.30	0.069	−0.17	0.210
354	−0.34	0.038	0.18	0.177	−0.25	0.097
500	−0.40	0.021	0.51	0.002	−0.00	0.494
707	−0.52	0.001	0.59	0.000	−0.08	0.332
1,000	−0.57	0.000	0.42	0.006	−0.30	0.044
1,414	−0.65	0.000	0.65	0.000	−0.18	0.190
2,000	−0.64	0.000	0.51	0.001	−0.30	0.051
22,050	−0.51	0.002	0.34	0.034	−0.30	0.058

Note that the same correlation coefficients can be associated with different *P* values because, as mentioned in the methods section, *P* values were computed using a statistical resampling technique (bootstrap), which took into account the correlation between left- and right-ear data, separately for each variable

loss duration, were observed for all CFs except the lowest two, i.e., 250 and 354 Hz, where PC was generally low and close to chance. The correlation coefficients are listed in Table 2 with the corresponding *P* values, and their magnitudes are illustrated in Figure 6, with the corresponding 95% confidence intervals. Statistically significant correlations are indicated by lower error bars terminating above zero.

The last two columns in Table 2 list correlation coefficients for PC and age, with the corresponding *P* values. As expected from previous studies showing decreases in speech-recognition performance with aging (for a review, see e.g., Gordon-Salant 2005), these correlations coefficients were generally negative. In fact, all eight correlation coefficients between PC and age were negative—which, according to a sign test, had a probability of $P=0.0039$ of occurring by chance. However, it is important to note that these correlation coefficients for PC and age were generally

smaller in magnitude than the correlations coefficients for PC and hearing-loss onset age or PC and hearing-loss duration, and unlike the latter, they only reached statistical significance ($P<0.05$) for one of the CF conditions (CF=1,000 Hz). The lack of significant correlation between PC and age for most of the CF conditions tested in the current study can be explained by considering that the variable, age, was equal to the sum of hearing-loss onset age and hearing-loss duration, and that the *negative* correlations between PC and hearing-loss onset age that were observed in several CF conditions approximately canceled out the *positive* correlations between PC and hearing-loss duration that were observed in the same conditions.

The two upper panels of Figure 7 shows scatter plots of PC versus hearing-loss onset age (upper left panel) and PC versus hearing-loss duration (upper right panel) for the 1,414-Hz CF condition. This condition was chosen for illustration purposes, as it yielded the largest correlation coefficients (in absolute value). The dashed lines are regression lines. For both hearing-loss onset age and hearing-loss duration, the linear regression explained approximately 42% of the variance in PC across ears. For comparison, the bottom panel in Figure 7 illustrates the relationship between PC and AI', for the same CF (1,414 Hz). The corresponding R^2 was equal to 19%. This indicates that, for the HI subjects tested in this study, hearing-loss onset age and hearing-loss duration were more informative than AI' for predicting speech-recognition performance. While these observations are based on the results of the 1,414-Hz CF condition, qualitatively similar conclusions were reached based on the results obtained for other CFs—except the lowest two, where the PC values were usually close to chance and not sufficiently variable to yield meaningful statistical results.

Since neither hearing-loss onset age nor hearing-loss duration were significantly correlated with pure-tone thresholds, speech levels, or AI' for any of the

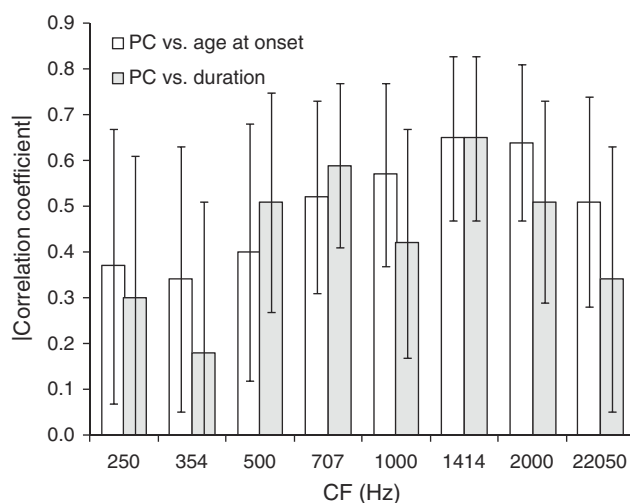


FIG. 6. Magnitudes of the correlation coefficients for PC and age at hearing-loss onset and PC and hearing-loss duration. The former are shown as empty bars; the latter are shown as filled bars. The error bars show 95% (bootstrap) confidence intervals.

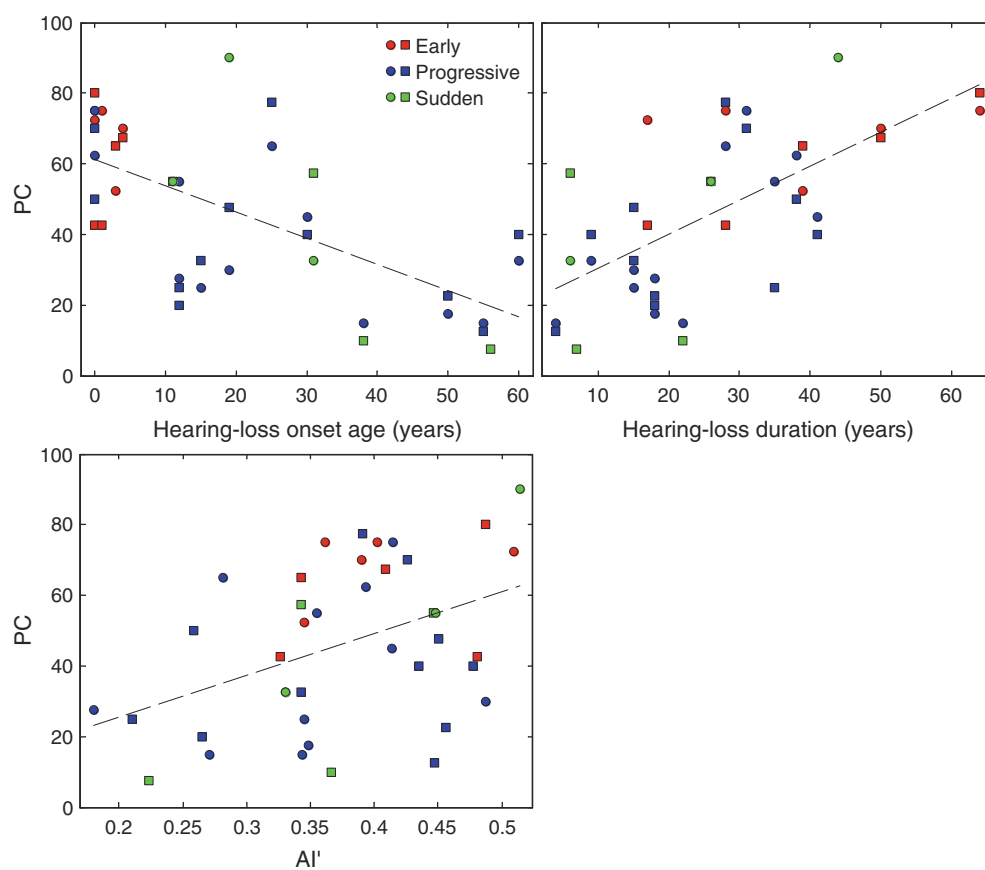


FIG. 7. Scatter plots of PC in the 1,414-Hz CF condition versus hearing-loss onset age, hearing loss duration, and AI' across all HI ears. Each data point corresponds to a single ear, in a single subject. *Circles* indicate the right ear, *squares* indicate the left ear, and the different colors indicate different hearing-loss onset groups, using the same color-group scheme as in previous figures. The *dashed lines* show regression lines through the data points.

CF conditions tested, the above-noted correlations between PC and hearing-loss onset age or hearing-loss duration cannot be explained (statistically) by correlations between PC and any of the following variables: AI', pure-tone thresholds, or speech levels. This was verified by computing correlations between PC and hearing-loss onset age, and between PC and hearing-loss duration, with and without the influence of AI', pure-tone thresholds, or speech levels partialled out. With one exception (for the CF=22,050 Hz condition, for which the correlation between hearing-loss duration and PC became nonsignificant when AI' was partialled out), correlations that were statistically significant when AI', pure-tone thresholds, and speech levels were not partialled out remained significant when any of these variables was partialled out.

The interpretation of simple correlations between PC and either hearing-loss onset age or hearing-loss duration, which are illustrated in Figure 6, is complicated by the existence of correlations between the latter two variables—expectedly, since hearing-loss duration was obtained by subtracting hearing-loss onset age from age. Significant correlations between

PC and hearing-loss onset age could be by-products of correlations between PC and hearing-loss duration, or the other way around. To determine whether the correlations with PC were driven by hearing-loss onset age, hearing-loss duration, or both, we computed partial-correlation coefficients between PC and each of the two variables, with the influence of the other variable partialled out. Figure 8 shows the magnitude of the resulting partial-correlation coefficients for PC and hearing-loss onset age, with the effect of hearing-loss duration partialled out (empty bars), and for PC and hearing-loss duration, with the effect of hearing-loss onset partialled out (filled bars). The partial-correlation coefficients are listed in Table 3. Significant negative partial correlations between PC and hearing-loss onset age were found for CFs higher than 707 Hz; for lower CFs, the partial correlations between these two variables were not significantly different from zero. Significant positive partial correlations between PC and hearing-loss duration were observed for CFs of 500, 707, and 1,414 Hz; for the other CFs, the correlations were not statistically significant. Although the above statistical-significance statements are based on a two-tailed critical probability level of

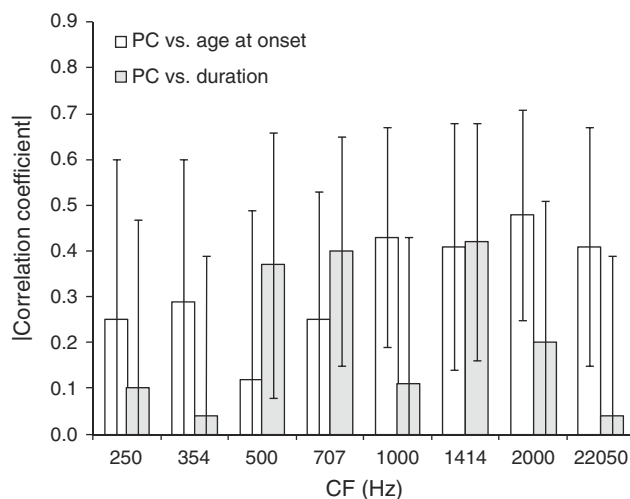


FIG. 8. Magnitudes of the partial correlation coefficients for PC and age at hearing-loss onset, and PC and hearing-loss duration, with the effect of the remaining variable partialled out. The *empty bars* correspond to partial correlation coefficients between PC and age at hearing-loss onset with the effect of hearing-loss duration partialled out. The *filled bars* correspond to partial correlation coefficients between PC and hearing-loss duration with the effect of age at hearing-loss onset partialled out. The *error bars* show 95% (bootstrap) confidence intervals.

0.05 per comparison, we determined (using bootstrap and taking into account the correlations in PC values across CFs) that the probability of finding three or more significant correlation coefficients out of eight by chance was slightly lower than 0.05. Since at least three significant correlations were observed for each variable (i.e., age at hearing loss onset and hearing-loss duration), the null hypothesis can be rejected.

In the above-described analyses of correlations, the relationship between PC and other variables was considered for each CF separately. However, an interesting question, which these analyses do not address, is whether the dependence of PC on CF was related to hearing-loss onset age or duration; otherwise stated: was the rate at which PC increased with CF correlated with

hearing-loss onset age or duration? To determine the answer to this question, while at the same time taking into account the influence of audibility, the following analysis was performed. Firstly, the slopes of individual “articulation functions” (i.e., PC as a function of AI’) were computed. For determining the slope, only CFs over which AI’ increased with CF were included. Secondly, correlation coefficients between these slopes and either hearing-loss onset age or hearing-loss duration were computed. The resulting correlation coefficients were equal to $r=-0.38$, with $P=0.008$ for slope versus onset age, and $r=0.48$, with $P=0.001$ for slope versus hearing-loss duration. The correlation between articulation-function slope and hearing-loss onset age was no longer significant after partialing out the influence of hearing-loss duration ($r=-0.12$ and $P=0.246$); however, the correlation between articulation-function slope and hearing-loss duration remained significant after hearing-loss onset age was partialled out ($r=0.34$ and $P=0.017$).

Finally, we examined correlations between PC and the duration of hearing-aid use. Because the duration of hearing-aid use was significantly correlated with hearing loss duration ($r=0.46$ and $P=0.017$), and the latter variable was found earlier to be significantly correlated with PC, we expected significant correlations also between hearing-aid use duration and PC. In fact, significant correlations between these two variables were only observed for two of the eight CF conditions tested (CF=1,414 Hz: $r=0.33$ and $P=0.042$ and CF=22,050 Hz: $r=0.34$ and $P=0.035$). When hearing-loss duration was partialled out, no significant partial correlation was found between the duration of hearing-aid use and PC.

DISCUSSION

The main findings of this study can be summarized as follows. Firstly, HI subjects with early-onset hearing

TABLE 3

Partial correlation coefficients for PC and hearing-loss onset age with hearing-loss duration partialled out, and PC and hearing-loss duration with hearing-loss onset age partialled out, and corresponding P values

CF	Onset age		Duration	
	r_{partial}	P	r_{partial}	P
250	-0.25	0.119	0.10	0.326
354	-0.29	0.065	-0.04	0.413
500	-0.12	0.284	0.37	0.027
707	-0.25	0.082	0.40	0.008
1,000	-0.43	0.006	0.11	0.279
1,414	-0.41	0.016	0.42	0.014
2,000	-0.48	0.003	0.20	0.142
22,050	-0.41	0.016	0.04	0.423

Note that the same correlation coefficients can be associated with different P values because, as mentioned in “Methods”, P values were computed using a statistical resampling technique (bootstrap), which took into account the correlation between left- and right-ear data, separately for each variable

loss (defined as hearing loss incurred before the age of four, with no major change thereafter) had higher mean speech-recognition scores than subjects with later, progressive or sudden hearing loss; the difference was statistically significant for CFs of 1 kHz and higher. Secondly, across all HI subjects, and in several of the CF conditions tested, speech-recognition scores were found to be negatively correlated with hearing-loss onset age, and positively correlated with hearing-loss duration. In the following, we consider possible explanations for these findings.

Higher speech intelligibility in subjects with early-onset hearing loss and correlation between speech intelligibility and hearing-loss onset age

Since the three HI groups had similar hearing thresholds, speech-presentation levels, and speech audibility (as measured by AI') on average, and similar age characteristics, the finding of higher speech-recognition scores in the "early" group than in the "progressive" and "sudden" groups cannot be explained simply by differences in audibility or age. One variable that differed significantly between three groups was hearing-loss onset age, which was lower (by definition) for the "early" group than for the two other groups. The groups also differed with respect to hearing-loss duration, which was longer for the "early" group—a by-product of the negative relationship between hearing-loss duration and hearing-loss onset age. When the effect of hearing-loss duration was partialled out—by entering it as a covariate in the ANOVA—the three HI groups were no longer found to differ significantly with respect to speech-recognition scores. One interpretation of this outcome is that the difference in speech-recognition scores between the "early" group and the other two groups was due entirely to a difference in hearing-loss duration between the groups. However, since hearing-loss duration co-varied with hearing-loss onset age, including one of these two variables as cofactor necessarily reduced the proportion of variance in speech-recognition scores that was left for the other variable to explain. Therefore, the absence of a significant group effect in the ANCOVA with hearing-loss duration as a cofactor should not necessarily be interpreted as evidence that the effect was driven entirely by differences in hearing-loss duration. Further study is needed to determine whether speech-recognition performance is significantly higher on average in subjects with "early" hearing loss onset than in subjects with later onset *even after hearing-loss duration is equated*. What the current results do demonstrate, on the other hand, is a relationship between speech intelligibility and hearing-loss onset age *even after the*

influence of hearing-loss duration is partialled out. This evidence comes from the finding of significant partial correlations between PC and hearing-loss onset age, in which correlations between each of these variables and hearing-loss duration were partialled out. Such significant partial correlations were observed for the four highest CFs (1, 1.41, 2, and 22.05 kHz).

In the current state of knowledge, we can only speculate as to the origin of these correlations between speech intelligibility and hearing-loss onset age, and as to why they were only observed for CFs of 1 kHz or higher. A possible explanation is in terms of neural plasticity, or perceptual learning. According to this explanation, the earlier the hearing loss, the more likely it is that subjects will learn—or that their central auditory systems will develop the neural circuits that are needed—to successfully decode the impoverished and distorted speech signals that they receive through their damaged peripheral auditory system. In this context, the finding of significant partial correlations between hearing-loss onset age and speech intelligibility for CFs of 1 kHz or higher can be explained by considering that at these frequencies, all of the ears tested had severe (≥ 70 dB HL) or profound (≥ 90 dB HL) hearing loss. As a result, speech cues in this frequency region were probably minimal, and strongly distorted. Being able to take advantage of such minimal cues to supplement speech-recognition performance—beyond the level that could be achieved using lower-frequency, less-distorted information only—may require special perceptual abilities; it is possible that individuals who were relatively young and had a highly plastic auditory system at the time at which they incurred a hearing loss are more able (or more likely) to develop such abilities than individuals in whom hearing loss was incurred later in life.

While plasticity provides one possible explanation for the findings of higher speech-recognition performance in subjects with early-onset hearing loss and of negative correlations between speech-recognition performance and hearing loss onset age, alternative explanations cannot be ruled out at this stage. For instance, early and late hearing losses may involve different patterns of inner and outer hair-cell damage or different degrees of auditory-nerve degeneration. This could lead to different impact on neural representations of speech in early and late loss, on average, without necessarily being reflected in different pure-tone thresholds. Alternatively, or in addition, because hearing generally deteriorates with age, and the subjects of the three hearing-loss onset groups were selected to have approximately the same amount of hearing loss on average than the subjects of the other two groups, it is possible that the selection

criteria that were used in this study created a bias, whereby subjects of the early-onset group tended to have lighter insults initially than subjects in the other two groups. Unfortunately, we did not have enough information concerning hearing-loss etiology and cochlear function in these subjects to address in detail the question of relationships between etiology and speech-understanding ability. This important question warrants further investigation.

Correlations between speech intelligibility and hearing-loss duration

The finding of significant correlations between hearing-loss duration and speech intelligibility was not completely unexpected. Previous studies in cochlear-implant listeners have shown that the duration of hearing-loss prior to cochlear implantation was a strong predictor of postimplantation performance in speech-perception tasks (Gantz et al. 1993; Blamey et al. 1996; Rubinstein et al. 1999). However, in these earlier studies, the correlation between hearing-loss duration and speech-recognition performance was *negative*: longer durations of hearing loss were associated with *worse* performance. In contrast, the results of the current study show a *positive* relationship between hearing loss duration and performance. How can these two apparently discrepant sets of results be reconciled? One possible explanation is that, while in the cochlear-implant studies duration of deafness represents a time when the auditory system was probably nearly totally deprived of input, in the current study, duration of deafness is time spent with some auditory input. Another—more speculative—explanation, relates to the fact that cochlear implants, being inserted basally into the cochlea, stimulate primarily auditory-nerve fibers that normally respond to relatively high frequencies. Therefore, it is likely that cochlear-implant listeners rely primarily or exclusively on information conveyed by “high-frequency” fibers in order to understand speech. On the other hand, the listeners in the current study had severe hearing loss above 1 kHz. Therefore, they had to rely primarily or exclusively on relatively low frequencies for speech understanding. It is conceivable that those listeners who had a hearing-loss for many years progressively became more accustomed to relying solely or primarily on low-frequency information (conveyed by auditory neurons with low characteristic frequencies) than the listeners in whom the loss of high-frequency hearing was more recent. Similar ideas have been proposed by Rhodes (1966), and more recently, Vestergaard (2003) and Moore and Vinay (2009). The latter two publications reported that listeners with cochlear dead regions at high frequencies were able to make better use of low-

frequency information for speech understanding than listeners who did not have such dead regions. The listeners in the present study were not tested for dead regions. It would be interesting, in a future study, to test whether longer-duration hearing losses are more likely to be associated with dead regions than shorter-duration hearing losses, and whether this is the case even when the listeners in the two groups have similar audiograms.

An enhanced ability to use low-frequency auditory channels (stimulated acoustically) at the expense of high-frequency channels (stimulated electrically by cochlear implants) could reflect tonotopic-map reorganization in the central auditory system following cochlear damage. Specifically, neurophysiological studies have shown that in animals with high-frequency hearing loss induced by mechanical or chemical damage to the basal part of the cochlea, primary auditory cortex neurons that formerly responded to high frequencies—within the frequency region where a hearing loss is now present—shift their preferred frequency downward—usually, within a frequency region that corresponds to the “edge” of the cochlear lesion (Rajan et al. 1993; for reviews, see: Irvine and Rajan 1996; Irvine et al. 2001). We can only speculate as to whether or not such central reorganization occurred in our human listeners and, if it did, whether it had significant functional consequences on speech understanding. To the best of our knowledge, the only perceptual effects that have been identified as possible consequences of tonotopic-map reorganization following cochlear damage so far are relatively small improvements in frequency-discrimination thresholds localized near the hearing-loss edge (McDermott et al. 1998; for a review, see Thai-Van et al. 2007). However, it is conceivable that central neural plasticity, which may develop over the course of several years following cochlear damage, has significant consequences on the processing of complex stimuli such as speech, which have not yet been identified. In fact, part of the large variability in speech-perception performance that has been observed across hearing-impaired individuals in this study and previous ones might stem from differences in the extent and functional success of central neural reorganization following peripheral damage.

Finally, it is worth noting that some studies have reported improvements in speech-recognition scores related to hearing-aid use, even when these scores were measured without the hearing aid—an effect known as “acclimatization” (for reviews, see Palmer et al. 1998; Turner et al. 1996; see also the note by Turner and Bentler 1998). Because 17 of the 20 subjects who took part in this study wore hearing aids on a regular basis (three monaurally, the other 14, binaurally), and the duration of hearing-aid use tended to increase with the

duration of hearing loss, it is conceivable that the positive correlations between hearing-loss and speech-recognition performance that were observed in the current study were actually driven by long-term acclimatization effects. Since acclimatization effects have rarely been documented over time periods exceeding a few months to a year, we do not know whether such effects can account for slow improvements in speech intelligibility over the course of several decades. To gain some clarity on this issue, one could do a study similar to the current one in a group of patients with varied (long as well as short) hearing-loss durations, who have never worn hearing aids.

Although the present results do not allow us to determine whether, and why, speech-recognition performance in HI subjects depends (in a causal way) on hearing-loss onset age and hearing-loss duration, these results demonstrate that these two variables can account for part of the large interindividual variability in speech-recognition performance among HI individuals with severe-to-profound high-frequency hearing loss.

ACKNOWLEDGMENTS

This work was supported by Vibrant Med-El France (Doctoral Research Grant CIFRE 266/2007 to F.S.) and the French National Center for Scientific Research (CNRS). Associate Editor Robert Carlyon and three anonymous reviewers greatly helped to improve the manuscript by suggesting important supplementary analyses, including the AI calculations. Special thanks are due to Liz Anderson and Christopher Turner, who provided detailed feedback on an earlier version of the manuscript. Michel Beliaeff, Christian Berger-Vachon, Sami Labassi, Peggy Nelson, Bénédicte Philibert, and Bert Schlauch are acknowledged for their support and helpful suggestions.

APPENDIX

To fit the psychometric functions relating PC to CF in the groups of NH and HI subjects, we used the following mathematical function, which is based on the equal-variance Gaussian signal-detection-theory model for the 1-of- m identification task (Green and Dai 1991; Macmillan and Creelman 2005).

$$P_{\text{mafc}}(\text{CF}) = 100 \left[\frac{\lambda}{m} + (1 - \lambda) \int_{-\infty}^{+\infty} \phi(z - d'(\text{CF})) \Phi^{m-1}(z) dz \right] \quad (6)$$

In this equation, $P_{\text{mafc}}(\text{CF})$ is the predicted PC for a given CF; λ is the asymptotic error rate; ϕ

denotes the standard normal function; Φ denotes the cumulative standard function; m is the number of response alternatives (which was set to 288 here, reflecting all possible pairwise combinations of 16 vowels and 18 consonants); and d' is the index of sensitivity. A power-law relationship between d' and CF was assumed: $d'(\text{CF}) = \text{CF}^\alpha / \sigma$, where σ is the standard deviation of the internal noise. The function described in Eq. 6 was used to fit the PC data using a least-squares approach (implemented with Matlab's *fminsearch* function, which uses the Nelder–Mead algorithm for function minimization).

REFERENCES

- ANSI (1969) American National Standard methods for the calculation of the speech intelligibility index, ANSI S3.5-1969. ANSI, New York
- ANSI (1997) American National Standard methods for the calculation of the speech intelligibility index, ANSI S3.5-1997. ANSI, New York
- BAER T, MOORE BC, KLUK K (2002) Effects of low pass filtering on the intelligibility of speech in noise for people with and without dead regions at high frequencies. *J Acoust Soc Am* 112:1133–1144
- BLAMEY P, ARNDT P, BERGERON F, BREDBERG G, BRIMACOMBE J, FACER G, LARKY J, LINDSTROM B, NEDZELSKI J, PETERSON A, SHIPP D, STALLER S, WHITFORD L (1996) Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants. *Audiol Neurotol* 1:293–306
- BUSBY PA, TONG YC, CLARK GM (1992) Psychophysical studies using a multiple-electrode cochlear implant in patients who were deafened early in life. *Audiology* 31:95–111
- BUSBY PA, TONG YC, CLARK GM (1993) Electrode position, repetition rate, and speech perception by early- and late-deafened cochlear implant patients. *J Acoust Soc Am* 93:1058–1067
- CHING TY, DILLON H, BYRNE D (1998) Speech recognition of hearing-impaired listeners: predictions from audibility and the limited role of high-frequency amplification. *J Acoust Soc Am* 103:1128–1140
- DAWSON PW, BLAMEY PJ, ROWLAND LC, DETTMAN SJ, CLARK GM, BUSBY PA, BROWN AM, DOWELL RC, RICKARDS FW (1992) Cochlear implants in children, adolescents, and prelinguistically deafened adults: speech perception. *J Speech Hear Res* 35:401–417
- DUBNO JR, DIRKS DD, ELLISON DE (1989a) Stop-consonant recognition for normal-hearing listeners and listeners with high-frequency hearing loss. I: the contribution of selected frequency regions. *J Acoust Soc Am* 85:347–354
- DUBNO JR, DIRKS DD, SCHAEFER AB (1989b) Stop-consonant recognition for normal-hearing listeners and listeners with high-frequency hearing loss. II: articulation index predictions. *J Acoust Soc Am* 85:355–364
- EFRON B, TIBSHIRANI R (1994) An introduction to the bootstrap. Chapman & Hall, New York
- FISHER RA (1915) FREQUENCY DISTRIBUTION OF THE VALUES OF THE CORRELATION COEFFICIENT IN SAMPLES OF AN INDEFINITELY LARGE POPULATION. *Biometrika* 10:507–521
- FISHER RA (1924) The distribution of the partial correlation coefficient. *Metron* 3:329–332
- FOURNIER JE (1951) *Audiométrie vocale*. Maloine, Paris
- FRENCH NR, STEINBERG JC (1947) Factors governing the intelligibility of speech sounds. *J Acoust Soc Am* 19:90–119

- FRIEDLAND DR, VENICK HS, NIPARKO JK (2003) Choice of ear for cochlear implantation: the effect of history and residual hearing on predicted postoperative performance. *Otol Neurotol* 24:582–589
- GANTZ BJ, WOODWORTH GG, KNUTSON JF, ABBAS PJ, TYLER RS (1993) Multivariate predictors of audiological success with multichannel cochlear implants. *Ann Otol Rhinol Laryngol* 102:909–916
- GANTZ BJ, HANSEN MR, TURNER CW, OLESON JJ, REISS LA, PARKINSON AJ (2009) Hybrid 10 clinical trial: preliminary results. *Audiol Neurotol* 14(Suppl 1):32–38
- GORDON-SALANT S (2005) Hearing loss and aging: new research findings and clinical implications. *J Rehabil Res Dev* 42(Suppl 2):9–24
- GREEN DM, DAI HP (1991) Probability of being correct with 1 of M orthogonal signals. *Percept Psychophys* 49:100–101
- GREEN KMJ, BHATT YM, MAWMAN DJ, O'DRISCOLL MP, SAEED SR, RAMSDEN RT, GREEN MW (2007) Predictors of audiological outcome following cochlear implantation in adults. *Cochlear Implants Int* 8:1–11
- HINDERINK JB, MENS LH, BROKX JP, VAN DEN BROEK P (1995) Performance of prelingually and postlingually deaf patients using single-channel or multichannel cochlear implants. *Laryngoscope* 105:618–622
- HOGAN CA, TURNER CW (1998) High-frequency audibility: benefits for hearing-impaired listeners. *J Acoust Soc Am* 104:432–441
- IRVINE DR, RAJAN R (1996) Injury- and use-related plasticity in the primary sensory cortex of adult mammals: possible relationship to perceptual learning. *Clin Exp Pharmacol Physiol* 23:939–947
- IRVINE DR, RAJAN R, BROWN M (2001) Injury- and use-related plasticity in adult auditory cortex. *Audiol Neurotol* 6:192–195
- KAMM CA, DIRKS DD, BELL TS (1985) Speech recognition and the Articulation Index for normal and hearing-impaired listeners. *J Acoust Soc Am* 77:281–288
- KESSLER DK, LOEB GE, BARKER MJ (1995) Distribution of speech recognition results with the Clarion cochlear prosthesis. *Ann Otol Rhinol Laryngol* 166:283–285
- KUJAWA SG, LIBERMAN MC (2009) Adding insult to injury: cochlear nerve degeneration after “temporary” noise-induced hearing loss. *J Neurosci* 29:14077–14085
- LEUNG J, WANG NY, YEAGLE JD, CHINNICI J, BOWDITCH S, FRANCIS HW, NIPARKO JK (2005) Predictive models for cochlear implantation in elderly candidates. *Arch Otolaryngol Head Neck Surg* 131:1049–1054
- MACMILLAN NA, CREELMAN CD (2005) Detection theory: a user's guide. Erlbaum, Mahwah
- MCDERMOTT HJ, LECH M, KORNBLUM MS, IRVINE DRF (1998) Loudness perception and frequency discrimination in subjects with steeply sloping hearing loss: possible correlates of neural plasticity. *J Acoust Soc Am* 104:2314–2325
- MOORE BCJ, VINAY SN (2009) Enhanced discrimination of low-frequency sounds for subjects with high-frequency dead regions. *Brain* 132:524–536
- MOORE BCJ, GLASBERG B, SCHLUETER A (2010) Detection of dead regions in the cochlea: relevance for combined electric and acoustic stimulation. *Adv Otorhinolaryngol* 67:43–50
- MURRAY N, BYRNE D (1986) Performance of hearing-impaired and normal hearing listeners with various high-frequency cut-offs in hearing aids. *Aust J Audiol* 8:21–28
- NIEMEYER W (1972) Studies on speech perception in dissociated hearing loss. In: Fant G (ed) International symposium on speech communication ability and profound deafness. The Alexander Graham Bell Association for the Deaf, Inc. pp 106–118
- OKAZAWA H, NAITO Y, YONEKURA Y, SADATO N, HIRANO S, NISHIZAWA S, MAGATA Y, ISHIZU K, TAMAKI N, HONJO I, KONISHI J (1996) Cochlear implant efficiency in pre- and postlingually deaf subjects. A study with H₂¹⁵O and PET. *Brain* 119:1297–1306
- PALMER CV, NELSON CT, LINDLEY GA (1998) The functionally and physiologically plastic adult auditory system. *J Acoust Soc Am* 103:1705–1721
- PAVLOVIC CV (1984) Use of the articulation index for assessing residual auditory function in listeners with sensorineural hearing impairment. *J Acoust Soc Am* 75:1253–1258
- RAJAN R, IRVINE DR, WISE LZ, HEIL P (1993) Effect of unilateral partial cochlear lesions in adult cats on the representation of lesioned and unlesioned cochleas in primary auditory cortex. *J Comp Neurol* 338:17–49
- RANKOVIC CM (1991) An application of the articulation index to hearing aid fitting. *J Speech Hear Res* 34:391–402
- RANKOVIC CM (1998) Factors governing speech reception benefits of adaptive linear filtering for listeners with sensorineural hearing loss. *J Acoust Soc Am* 103:1043–1057
- RANKOVIC CM (2002) Articulation index predictions for hearing-impaired listeners with and without cochlear dead regions. *J Acoust Soc Am* 111:2545–2548
- RHODES RC (1966) Discrimination of filtered CNC lists by normals and hypacusics. *J Aud Res* 6:129–133
- RUBINSTEIN JT, PARKINSON WS, TYLER RS, GANTZ BJ (1999) Residual speech recognition and cochlear implant performance: effects of implantation criteria. *Am J Otol* 20:445–452
- SHIPP DB, NEDZELSKI J (1995) Prognostic indicators of speech recognition performance in adult cochlear implant users: a prospective analysis. *Ann Otol Rhinol Laryngol Suppl* 166:194–196
- THAI-VAN H, MICHEYL C, NORENA A, VEUILLET E, GABRIEL D, COLLET L (2007) Enhanced frequency discrimination in hearing-impaired individuals: a review of perceptual correlates of central neural plasticity induced by cochlear damage. *Hear Res* 233:14–22
- TONG YC, BUSBY PA, CLARK GM (1988) Perceptual studies on cochlear implant patients with early onset of profound hearing impairment prior to normal development of auditory, speech, and language skills. *J Acoust Soc Am* 84:951–962
- TURNER CW, BENTLER RA (1998) Does hearing aid benefit increase over time? *J Acoust Soc Am* 104:3673–3674
- TURNER CW, HUMES LE, BENTLER RA, COX RM (1996) A review of the literature on hearing aid benefit as a function of time. *Ear Hear* 17:14S–28S
- VAN DIJK JE, VAN OLPHEEN AF, LANGEREIS MC, MENS LHM, BROKX JPL, SMOORENBURG GF (1999) Predictors of cochlear implant performance. *Audiology* 38:109–116
- VESTERGAARD MD (2003) Dead regions in the cochlea: implications for speech recognition and applicability of articulation index theory. *Int J Audiol* 42:249–261
- VICKERS DA, MOORE BC, BAER T (2001) Effects of low-pass filtering on the intelligibility of speech in quiet for people with and without dead regions at high frequencies. *J Acoust Soc Am* 110:1164–1175

II.3. Reconnaissance de la parole filtrée passe-bas chez les sujets ayant une surdité dans les hautes fréquences

(Article soumis à J Assoc Res Otolaryngol. : Seldran F, Michey C, Thai-Van H, Truy E, Gallego S. Lowpass-filtered speech recognition in ears with high-frequency hearing loss)

Dans les deux premières études nous avons constaté que chez les malentendants avec des surdités dans les hautes fréquences, il existait de larges disparités dans les fonctions psychométriques « Intelligibilité = f (fréquence de coupure de la parole filtrée passe-bas) ». Trois causes identifiées de cette variabilité sont l'audibilité mais aussi et surtout la durée de surdité et l'âge d'apparition de la surdité.

Cette troisième étude a consisté à classifier les patients en catégories, à l'aide d'un algorithme de regroupement (clustering) basé sur les audiogrammes des patients. Quarante-trois sujets malentendants ont participé à cette étude, ce qui nous a permis de travailler sur un pool de 86 oreilles avec différentes configurations d'audiogrammes qui variaient en terme de pente et de perte auditive dans les basses fréquences.

Pour nous assurer que les résultats confirmaient ce que nous avons trouvé dans notre article 7, nous avons cherché si les performances étaient toujours reliées à la durée de privation auditive et à l'âge d'apparition de la surdité. Les résultats obtenus sur cette plus large population confirment ces relations.

Les fonctions psychométriques (intelligibilité en fonction de la fréquence de coupure) obtenues pour les différents groupes d'oreilles sont à peu près regroupées de la même façon que les audiogrammes. Ce qui signifie, comme on peut s'y attendre, que les groupes qui ont la meilleure audition résiduelle dans les basses fréquences ont les meilleures performances. Ceci confirme le résultat de l'étude précédente montrant que l'intelligibilité est en grande partie liée à l'audibilité.

Par ailleurs, en regardant les fonctions psychométriques des individus inclus dans chacun des groupes, nous constatons une très large variabilité des performances. Plus particulièrement, à l'intérieur d'un même groupe, nous retrouvons les mêmes patterns de réponses que nous présentions dans l'article 6 du présent manuscrit. Cela signifie que notre test d'audiométrie vocale filtrée passe-bas peut s'avérer très utile pour évaluer si un sujet peut bénéficier de l'EAS ou d'une technologie alternative type amplification conventionnelle, compression fréquentielle ou implantation cochléaire. Actuellement ce test demande beaucoup de temps (environ 2h par patient) et il semble être intéressant de le simplifier, peut-être ne réaliser que les mesures « filtré à 700 Hz » et « non filtré », pour qu'il soit pratiqué en routine clinique.

Lowpass-filtered speech recognition in ears with high-frequency hearing loss

Fabien Seldran,^{1,2,3,4*} Christophe Micheyl,⁵ Hung Thai-Van,^{1,2,3,6} Eric Truy,^{1,2,3,6}

Stéphane Gallego^{1,2,3}

¹*INSERM U1028, Lyon Neuroscience Research Center, PACS Team (Speech, Audiology, Communication, Health), Lyon, F-69000, France*

²*CNRS UMR5292, Lyon Neuroscience Research Center, PACS Team (Speech, Audiology, Communication, Health), Lyon, F-69000, France*

³*University Lyon 1, Lyon, F-69000, France*

⁴*Vibrant Med-El Hearing Technology GmbH, 400 Ave. Roumanille, BP 309, Sophia-Antipolis 06906, France*

⁵*Department of Psychology, University of Minnesota, Minneapolis, MN 55455, USA*

⁶*Audiology and ENT Department, Edouard Herriot Hospital, Lyon F-69437, France*

*Corresponding author: Fabien Seldran, INSERM U1028, CNRS UMR5292, Centre de Recherche en Neurosciences de Lyon, Equipe Audition, Hôpital Edouard Herriot, Pavillon U, Place d'Arsonval, F-69437 Lyon Cedex 03, France. Tel: +33.4.72.11.05.03 Fax: +33.4.72.11.05.04

Author emails: fseldran@yahoo.fr (F. Seldran), cmicheyl@umn.edu (C. Micheyl), hthaivan@gmail.com (H. Thai-Van), eric.truy@chu-lyon.fr (E. Truy), sgallego@hotmail.fr (S. Gallego)

Running title: Hearing loss and lowpass-speech recognition

ABSTRACT

Hearing-impaired individuals differ widely in their ability to understand speech. The factors responsible for this variability are not completely understood. Although audibility plays a primordial role, recent studies suggest that other factors also contribute. In this study, we analyzed pure-tone thresholds and psychometric functions (PFs) for the recognition of lowpass-filtered speech in 86 hearing-impaired (HI) ears (43 subjects). A statistical clustering algorithm was used to separate the ears into seven groups corresponding to different audiogram types, ranging from steeply sloping to quasi-flat. The result showed that while different audiogram groups usually had different mean PFs, ears with the same audiogram type could exhibit widely different PFs. In addition, while speech-recognition scores co-varied with audibility-based predictions, they were also significantly correlated with the age at hearing-loss onset and the number of years of hearing-loss. Finally, for signals filtered at low CFs (<1 kHz), HI ears with steeply sloping audiograms had higher recognition scores than NH ears. These results illustrate the limitations of audiogram- or audibility-based predictions of speech-recognition performance in HI individuals, and they provide further evidence that speech-recognition performance in HI individuals depends on other factors than audibility, including the number of years of experience relying primarily or exclusively on low-frequency information for speech understanding.

Keywords: hearing loss; thresholds; audiogram; speech intelligibility

INTRODUCTION

Among individuals with high-frequency hearing loss, some are able to benefit from high-frequency acoustic stimulation, but others are not (Ching et al., 1998; Hogan & Turner, 1998; Murray & Byrne, 1986; Skinner & Miller, 1983; Turner & Cummings, 1999; Vickers et al., 2001; Villchur, 1973). For the latter group, electro-acoustic stimulation (EAS), which involves a combination of low-frequency acoustic stimulation via a hearing aid (HA) and electric stimulation via a short-electrode cochlear implant (CI), may provide more benefit than a conventional HA (Gantz et al., 2009; Turner, 2006; Woodson et al., 2010). However, since the surgical insertion of an electrode array into the basal part of the cochlea subsequently limits acoustic stimulation to relatively low frequencies—usually, below 1 kHz (Gantz et al., 2009; Turner, 2006; Woodson et al., 2010), it is essential to assess how much a patient who appears, based on his/her audiogram, as a potential EAS candidate, actually benefits from acoustic stimulation at high frequencies, prior to implantation. To this aim, it may be necessary to measure speech-recognition performance using speech signals filtered below or above a cutoff frequency (CF), systematically, for different CFs. However, considering that such detailed measurements are time consuming and, sometimes, impractical, it would be highly desirable for clinicians to be able to infer how much benefit a patient derives from acoustic stimulation at frequencies above a certain CF, using only information collected during routine audiometric visits, such as pure-tone thresholds, the patient's age, and the patient's history.

Previous research has shown that the audibility of spectral components of the speech signal that are important for intelligibility are a primordial determinant of differences in speech-recognition performance between HI individuals (e.g., Dubno et al., 1989; Hogan & Turner, 1998; Rankovic, 1991; 1998; 2002). However, even after audibility has been taken into account, a

substantial residual proportion of the variance in speech-recognition scores remains unexplained (Baer et al., 2002; Ching et al., 1998; Dubno et al., 1989; Hogan & Turner, 1998; Kamm et al., 1985; Pavlovic, 1984), suggesting that additional factors are involved. In a recent study, Seldran et al. (2011) found that HI subjects in whom hearing loss was incurred before the age of four had significantly higher speech-recognition scores, on average, than other HI subjects, in whom absolute thresholds and predicted speech audibility were roughly the same, but in whom hearing loss was incurred later in life. Moreover, they found that across all HI listeners, speech-recognition scores were negatively correlated with age at hearing-loss onset, and positively correlated with the duration of hearing loss. These findings indicate that other factors than audibility influence speech recognition performance in HI listeners.

In this context, the current study had two main goals. First, to investigate whether the way in which speech-recognition performance depends on the lowpass-filter cutoff frequency (CF) in HI ears can be predicted, based on the audiogram type, e.g., steeply sloping versus shallow sloping. Secondly, to follow-up on Seldran et al.'s (2011) study, and check whether these authors' findings of significant correlations between speech-recognition performance, age at hearing loss onset, and hearing-loss duration would hold in a larger sample of HI subjects, who have a wider variety of audiometric configurations, and more varied histories than in the previous study. In addition, relationships between speech-recognition scores, age at testing, and the duration of hearing-aid (HA) use were also tested.

METHODS

Listeners

Data from 86 HI ears (43 HI subjects; 27 female, 16 male, aged 21 to 74 years; median age = 56

years) and 20 normal hearing (NH) ears (20 subjects; 7 female, and 13 male, aged 20 to 32 years, median age = 23 years) were analyzed. Some of these data, including pure-tone thresholds and speech-recognition scores from 14 of the 43 HI subjects and from all 20 NH listeners, were collected in previous study, and have already been reported elsewhere (Seldran et al., 2011). For the purpose of the current study, these data were pooled with the data of 29 new HI subjects (58 new HI ears) to obtain a larger sample to allow more robust statistical inferences and more general conclusions than could be obtained using the smaller sample.

Table 1 lists the subject's gender and age, the duration of the hearing-loss (obtained by subtracting the age at hearing-loss onset from the age), the number of years of HA use, and the pure-tone threshold averages (PTAs), separately for each of the 86 HI ears. The stars in the first column indicate subjects who were tested in the previous study (Seldran et al., 2011). Although a total of 40 HI ears were tested in the previous study, data from only 28 of these ears were included in the present study in order to ensure that pure-tone threshold measurements in 1/4-octave steps were available for all of the ears included in the current study. The decision to measure absolute thresholds in 1/4-octave, rather than 1- or 1/2-octave steps, was motivated by the consideration that this would increase the precision of the statistical-clustering analyses and of the audibility-based predictions of speech intelligibility, both of which used these pure-tone thresholds (see below for details). The pure-tone threshold averages (PTAs) listed in Table 1 were computed by averaging pure-tone thresholds across all tested frequencies comprised between, and including, the indicated limits (250-850 Hz or 1-4 kHz). The last column in Table 1 indicates the known, or suspected, etiology of the hearing loss, when this information was available.

Table 1

Figure 1 shows the mean and individual audiograms (defined here as pure-tone thresholds at 1/4-octave steps between 250 and 4000 Hz) for the left and right HI ears. In the interest of time, and motivated by considerations of audibility and frequency-band importance considerations for speech recognition, pure-tone thresholds were only measured (in 1/4-octave steps) up to 4 kHz. Thresholds ranged from 20 dB HL or less, which was defined as “normal,” to higher than the maximum sound level that could be produced by the audiometer, which was defined as “unmeasurable”. This maximum level varied from 105 to 120 dB HL depending on the test frequency and channel, being equal to 120 dB HL for frequencies between and 600 and 2000 Hz. For data-visualization and data-analysis purposes, unmeasurable thresholds were set arbitrarily to a value of 125 dB HL. Out of the 86 HI ears tested, 25 had unmeasurable thresholds at 4 kHz. Air-bone gaps were never greater than 10 dB at any of the frequencies tested, indicating sensorineural, rather than conductive, hearing loss.

Figure 1

The study was approved by the Ethics Committee (CPP Sud-Est IV, Centre Leon Berard de Lyon, France, N° ID RCB: 2008-A01479-46). In accordance with the Declaration of Helsinki, written informed consent was obtained.

Stimuli and procedure

Speech-intelligibility measurements were performed using disyllabic French words (Fournier, 1951). These lists may be regarded as the French equivalent of the American “Spondees” lists. The words were spoken by a male speaker, and recorded on a Compact Disc (CD) at a 44.1-kHz sampling frequency with a 16-bit quantization range. The sampled signals were uploaded into the Adobe Audition software, which was used to lowpass-filter them digitally in the frequency

domain using the software's "FFT filter" function. This produced a lowpass filter with an abrupt transition band and a stopband attenuation of approximately 70 dB. All speech signals were lowpass-filtered separately at each of the following CFs: 250, 353, 500, 707, 1000, 1414, and 2000 Hz. These frequencies correspond to all half-octave steps between 250 and 2000 Hz. In addition to these seven "lowpass-filtered" conditions, an "unprocessed" condition was tested. In that condition, no lowpass-filtering was applied to the speech signals imported from the CD at 44.1 kHz. In the figures, the data for this condition are plotted as corresponding to a CF of 22050 Hz. For statistical data-analysis purposes, the data for this "full band" condition were not treated differently from the data obtained in the other (lowpass-filtered) test conditions. Finally, the processed and unprocessed signals were burned to CD (44.1-kHz sampling frequency, 16-bit quantization range), and presented monaurally to the listeners.

All eight processing conditions were tested in all listeners. For the HI listeners, both ears were tested, and 20 words (40 syllables) were presented at a single level in each test condition. The unprocessed (CF = 22050 Hz) condition was tested first, at three or four different sound levels: 70, 80, 90, and/or 100 dB SPL, where SPL refers to the root-mean-square (RMS) sound-pressure level of the stimuli across all stimuli. The highest level (100 dB SPL) was only tested in listeners who did not find it uncomfortably loud. In this initial test, and in all subsequent tests, the left and right ears of each subject were tested, separately, in random order. During this initial phase of the test, the subjects were asked to indicate which listening level they found most comfortable. These levels are listed in Table 1. Note that the comfortable level could differ between the left and right ears of a given subject. On average across all ears, the most comfortable level was equal to 95.35 dB SPL. Four subjects described the intensity of 90 dB SPL as too low, and that of 100 dB SPL as too loud. For these subjects, an intermediate level (95 dB

SPL) was used. The level that the subject indicated as comfortable was used during subsequent testing, which involved lowpass-filtered speech. However, because the bandwidth of the filtered stimuli decreased with CF, and was always smaller than for the “unprocessed” stimuli, the actual SPL of the filtered stimuli was lower than the SPL of the unprocessed (CF = 22050 Hz) stimuli, and it decreased with CF, as was the case in previous studies (e.g., Hogan & Turner, 1998; Vickers et al., 2001).

In NH listeners, eight speech levels, ranging from 10 to 80 dB SPL in steps of 10 dB, were tested for each of the CF conditions, including the “full band” condition. Because the resulting number of test conditions was considerably larger than for HI listeners, a single, randomly selected list of 10 words was presented in each test condition, and only the right ear was tested. The listener’s task was to repeat the words as they heard them.

Apparatus

Pure-tone audiometry was performed using a MADSEN Orbiter 922 audiometer. For the speech tests, the signals were played via PHILIPS CD723 CD player connected to the MADSEN audiometer. All sounds were presented monaurally via TDH39 earphones.

Data analysis

Quantification of speech-recognition performance

The number of syllables that was correctly repeated by the listener in a given test condition was divided by the total number of syllables presented in that condition (40 for HI listeners, 20 for NH listeners), and the result was multiplied by 100 to obtain a percent-correct (PC) score for that condition and that listener.

Audibility calculations and speech-intelligibility predictions

To investigate the influence of speech audibility on the speech-recognition performance of HI listeners, we used an approach similar to that described in the AI (ANSI S3.5, 1969). First, we computed the LTAS of the speech signals and expressed the spectrum levels into “speech sensation levels” using the following equation.

$$s_{i,j} = L_j + l_i - t_{i,j} - \theta_i, (1)$$

where $s_{i,j}$ is the speech “sensation level” corresponding to the i^{th} frequency band for the j^{th} HI listener; L_j is the speech presentation level (in dB SPL) used in that listener during the experiment; l_i is the LTAS level of speech (in dB SPL/Hz, assuming a 0-dB SPL overall presentation level) corresponding to frequency f_i ; $t_{i,j}$ is an estimate of the listener’s pure-tone threshold (in dB HL) at frequency f_i , which was computed using linear interpolation on the pure-tone thresholds measured in 1/4-octave steps in that listener; and θ_i is the “reference internal noise spectrum level” (in dB SPL/Hz) at frequency f_i , which was obtained using polynomial interpolation on the values given in the ANSI standard (ANSI S3.5, 1997). The frequencies, f_i ($i = 1, \dots, 8$), were set to: 125, 297, 420, 595, 841, 1189, 1682, and 4696 Hz. The middle six values in this list correspond to the geometric means of two consecutive lowpass-filter CFs; for example, 297 Hz is the geometric mean of 250 Hz (the lowest lowpass-filter CF used in this study) and 353 Hz (the next lowpass-filter CF value used in this study). The first and last values in the list correspond to the geometric mean of 62.5 Hz (i.e., one octave below 125 Hz) and 250 Hz, and to the geometric mean of 2000 Hz and 11025 Hz (i.e., one octave below the Nyquist frequency, 22050 Hz), respectively.

The $s_{i,j}$ ’s were used to compute a quantity analogous to the Articulation Index (ANSI

S3.5, 1969), hereafter referred to as $A_{k,j}$, using the following equation.

$$A_{k,j} = \sum_{i=1}^8 a_{i,j} w_i q_k(i), \quad (2)$$

with,

$$a_{i,j} = \max[\min(s_{i,j}, 30), 0] / 30 + 27, \quad (3)$$

and,

$$\begin{aligned} q_k(i) &= 1 \text{ if } i \leq k, \\ &= 0 \text{ if } i > k, \end{aligned} \quad (4)$$

where the subscript, k , refers to the k^{th} lowpass-filter CF, and the “importance” or “weight” of the i^{th} frequency band, w_i , was determined using the speech-recognition scores measured in the 20 NH listeners. Specifically, the vector of frequency-importance values, $\mathbf{w} = [w_1, \dots, w_8]^T$ (boldface characters are used to denote vectors, the superscript T denotes vector transposition) was computed by finding the least-squares solution to the matrix equation,

$$\mathbf{A}\mathbf{r} - \mathbf{w} = 0, \quad (5)$$

where the $b_{km,i}$ entry in the 64x8 matrix, \mathbf{A} , was computed as,

$$b_{km,i} = \sum_{j=1}^8 \alpha_{i,j} w_j, \quad (6)$$

with,

$$\alpha_{i,j} = \max[\min(10 + l_i - \theta_i, 30), 0] / 30 + c, \quad (7)$$

and each entry, r_k , in the vector, $\mathbf{r} = [r_1, \dots, r_8]^T$, was the rationalized arcsine transformation (Studebaker, 1985) of the mean PC across the 20 NH listeners for the k^{th} lowpass-filter CF. The rationalized-arcsine transformation is defined as,

$$r = \frac{146}{\pi} \left[\sin^{-1} \left(\sqrt{\frac{np/100}{n+1}} \right) + \sin^{-1} \left(\sqrt{\frac{np/100+1}{n+1}} \right) \right] - 23. \quad (8)$$

where r is the score in rationalized-arcsine units (raus), p is the PC, and n is the total number of presented syllables for the considered condition and listener.

The constant, c , in equation 7, was adjusted by recursively solving the following matrix equation using a least-squares criterion,

$$\mathbf{r} - \mathbf{A}\mathbf{w} = 0. \quad (9)$$

The resulting value of c , which was equal to 27, was plugged into equation 3.

The predicted speech-intelligibility scores ($A_{k,j}$'s) computed using this procedure were in raus. They were transformed back to PC scores using the inverse rationalized-arcsine transform,

$$p = 50 \left[1 - \left(\sqrt{\frac{n(n+2) - \cot^2\left(\frac{\pi(r+23)}{146}\right)}{n}} \right) \cos\left(\frac{\pi(r+23)}{146}\right) \right]. \quad (10)$$

Cluster analyses

One of the two approaches that we used to investigate relationships between speech-recognition scores and hearing thresholds involved organizing the data into groups. To this aim, we used a k -means clustering algorithm (Allen & Eddins, 2010; Duda et al., 2001; Lloyd, 1982; Steinhaus, 1956). One advantage of using a statistical clustering algorithm is that the algorithm sorts the data into groups based on a measure of similarity (inversely related to distance in the variable space) automatically, thus obviating the need for the researcher to define criteria based on potentially arbitrary considerations (Allen & Eddins, 2010). The only user-defined variable in the clustering algorithm was the number of groups, k . This number had to be specified beforehand, as it is not determined automatically by the k -means algorithm. There exist different approaches for the automatic determination of the number of clusters in datasets, all of which involve some

degree of arbitrariness in the definition of the criteria that are used to define the groups (Duda et al., 2001). Our primary objective, in this study, was to investigate relationships between pure-tone thresholds and speech-recognition scores, and not to determine how many groups could be distinguished in the HI-listeners' data. Based on visual exploration of the data, we determined that a k of 7 was large enough to capture the main types of pure-tone audiogram configurations that were apparent in the individual data, and small enough to ensure that each group contained data from at least 6 ears.

The k-means algorithm was applied to the pure-tone thresholds, resulting in the formation of seven groups of ears characterized by different audiograms. If pure-tone thresholds were a reliable indicator of speech-recognition performance, groups of ears that have different mean audiograms (e.g., flat versus steeply sloping audiograms, or steeply sloping versus more shallowly sloping audiograms) should also display different psychometric functions (PFs) for speech recognition (e.g., speech-recognition scores that increase continuously, saturate slowly, or saturate rapidly, as a function of CF).

Analyses of correlations between speech-recognition scores and other variables

To investigate statistical relationships between speech-recognition scores and other variables, including age, hearing-loss onset age, hearing-loss duration, and HA-use duration, we computed Pearson correlation coefficients. One complication in determining the statistical significance of these correlation coefficients stemmed from the lack of statistical independence between measures collected in the left and right ears of the same individual. Standard tests of statistical significance for correlation coefficients assume that the two variables being correlated are not internally correlated. To circumvent this problem, we estimated the sampling distribution of

Fisher-z-transformed correlation coefficients (Fisher, 1915) under the null hypothesis (i.e., no correlation) using a permutation technique (Efron & Tibshirani, 1993). Specifically, Fisher-z-transformed correlation coefficients for the two variables of interest were computed repeatedly (1000 times), with the ordering of the rows (corresponding to different listeners) shuffled on every trial, so that the pairing of the left and right ears was maintained at all times (Seldran et al., 2011). The maximum-likelihood estimate of the standard deviation of the Fisher-z-transformed values thus obtained was computed, using a Gaussian function—also known as “Laplace approximation.” This estimate was then used to determine a p value for the correlation coefficient, using the inverse-cumulative standard normal function. A similar procedure was used to determine p values for the difference between correlation coefficients for two pairs of variables.

RESULTS

Group-based analyses

Figure 2 shows the mean (left column) and individual (right column) audiograms for the seven groups of ears that were obtained using the clustering algorithm. The groups are organized into three panels, depending on the steepness of the audiogram slope: the three steep-slope groups in the upper row, the three moderate-slope groups in the middle row, and the quasi-flat-audiogram group in the bottom row. Table 2 lists the mean slopes for each group, along with additional information, including the mean drop in threshold between 250 and 1400 Hz. Within each panel, the different groups are identified by different colors: green for the steeper-slope group (in the upper and middle rows), blue for the intermediate-slope group (or, for the single group), and red for the shallower-slope group (in the upper and middle rows).

Figure 2

Table 2

The right-hand panels in Fig. 2 show that, although there was some variability in pure-tone thresholds within each group, there was relatively little overlap *across* the groups within each panel, at least, over the 250-1000 Hz frequency range. Moreover, over that range, the error bars, which show the 95% confidence intervals (bootstrap, $N = 10000$) of the group-mean thresholds, also do not overlap. This shows that, for the most part, the k-means algorithm was successful in forming separate groups based on pure-tone thresholds.

To facilitate comparisons between the audiograms of the two pairs of groups that had approximately the same mean pure-tone thresholds at 250 Hz across the two upper panels in Fig. 2, the mean and individual audiograms of these two pairs of groups have been re-plotted in Fig. 3. There, it is clearly apparent that the mean audiograms of the two groups in each pair had different slopes.

Figure 3

Figure 4 shows the mean and individual PFs for speech recognition for the same seven groups of ears, using the same arrangement of panels, symbols, and colors, as in Fig. 2. The dashed lines show the mean “normative” PFs that were measured in 20 NH listeners using the same speech stimuli, and presentation levels of 10-80 dB SPL, in 10-dB steps. The lowest three dashed curves show the mean PFs for levels of 10, 20, and 30 dB SPL (from low to high), while the fourth dashed curve shows the PF averaged across all other levels. Normative PFs measured using levels of 40 dB SPL and higher were averaged because, for the NH listeners, increases in presentation level above 40 dB SPL had no significant effect.

Figure 4

With the exception of group 6 (middle row, red downward-pointing triangles), in which pure-tone thresholds were larger than 90 dB HL on average even at 250 Hz, PC usually increased as a function of CF, at least, over some range of CFs. Moreover, with the exception of group 7 (bottom row), which had a quasi-flat mean audiogram, and for which no saturation in PC with increasing CF was apparent, for all other groups, PC usually reached a plateau for a CF lower than 22.05 kHz. Although the variability in the data makes it difficult to determine precisely the value of CF for which PC stopped increasing, visually, a “knee-point” in the PFs was usually apparent for CFs comprised between 500 and 2000 Hz.

The error bars, which show 95% confidence intervals (bootstrap, $N = 10000$) around the mean PCs, indicate that, except for CFs lower than 500 Hz, where PC was usually quite low (less than 20% on average), the mean speech-recognition scores of the different groups were often statistically distinguishable. Thus, for the most part, the different mean-audiogram configurations shown in Fig. 2 appear to correspond to different mean PFs. However, the 95% confidence intervals around the mean PCs were relatively large: typically, around 20 percentage points; this reflects the large variability in PFs across ears within each group, which is apparent in the right-hand panels of Fig. 4. For example, looking at the upper-right panel in Fig. 4, it can be seen that, in the group of the steep audiogram slopes, the PFs of at least 3 of the 19 ears in the intermediate-slope subgroup (blue circles) overlapped with the PFs of the steep-slope subgroup (green leftward-pointing triangles). Also, the PFs of at least 4 of the 11 ears in the shallow-slope subgroup (red rightward-pointing triangles) overlapped with the PFs of the intermediate-slope subgroup (blue circles). Even more substantial overlap is apparent between the PFs of the steep-slope subgroup (green upward-pointing triangles) and the intermediate-slope subgroup (blue squares) in the moderate-audiogram-slope group (middle-right panel). Here, the PFs of more

than one half of the ears in the two groups were overlapped across these subgroups. Such extensive overlap in PFs across these two subgroups would have been difficult to predict based on the audiograms measured in these same two subgroups, which showed relatively little overlap below 1000 Hz (Fig. 2).

Figure 5

The mean and individual PFs of the two pairs of subgroups that had similar mean thresholds at 250 Hz but different audiogram slopes (Fig. 3) have been re-plotted in Fig. 5 to facilitate direct visual comparisons between these subgroups. This figure shows that, given similar mean thresholds at 250 Hz, the shallower the audiogram slope, the steeper the PF slope and the higher the highest PC, on average. However, the figure also illustrates substantial variability across ears within each group, and substantial overlap across the groups.

Correlation analyses

In this section, we examine correlations between the measured speech-recognition scores and the following variables: the speech-recognition scores predicted based on audibility and importance functions; the listeners' age at testing; the age at hearing-loss onset; the duration of hearing loss; and the number of years of HA use at the time of testing, which is hereafter referred to as "duration of HA use," and abbreviated, "HA-use duration."

Correlations between measured and predicted speech-recognition scores

Figure 6 shows a scatter plot of the measured rau scores as a function of the scores predicted based on the audibility and importance function of the speech signals (as described in the

Methods, Data Analysis section). The Pearson correlation coefficient between these two variables was equal to 0.75 ($r^2 = 0.57$), and was statistically different from 0 ($p < 0.0001$). Note that this p value, and the ones listed in the tables below, were computed using the pseudo-random shuffling technique described in the Methods section; this yielded an estimate of the statistical distribution of p values, taking into account the influence of correlations between measures obtained in the left and right ears of the same subject.

Figure 6

One limitation of the preceding analysis is that speech-recognition scores corresponding to different CFs were pooled. Therefore, it is not clear, to what extent the relatively high r and r^2 values are due, simply, to the fact that the measured scores and the predicted scores both tended to increase as a function of CF. To determine whether the measured and predicted scores were still significantly correlated with each other after their common dependence on CF was removed, we computed correlation coefficients between measured and predicted scores for each CF, separately. These coefficients are listed in Table 3. For all except the lowest two CFs, these coefficients were positive, and significantly different from 0. Surprisingly, for the lowest CF (250 Hz), a significant correlation coefficient was observed. Visual inspection of the scatter plot between measured and predicted rau scores for this low CF revealed that the correlation was spurious, reflecting a floor effect—for over half of the HI ears tested, the PCs measured in that CF condition were equal to 0.

Table 3

Correlations between speech-recognition scores and other variables

Table 4 lists Pearson correlation coefficients between the rau scores measured in HI ears and the following variables: age at testing, age at hearing-loss onset, hearing-loss duration, and HA-use

duration. The subject's age at the time of testing was not significantly correlated (at the Bonferroni-corrected level: $p = 0.05/8 = 0.00625$) with speech-recognition scores for any of the CFs tested. In contrast, statistically significant correlations between hearing-loss onset age and speech-recognition scores were observed for all tested CFs except the lowest one (250 Hz). Statistically significant correlations were also observed for speech-recognition scores and hearing-loss duration, but only for CFs comprised between 500 Hz and 2000 Hz. Although the latter correlations coefficients were generally smaller than those involving hearing-loss onset age, the differences in the magnitudes of these coefficients were relatively small and not statistically significant ($0.12 < p < 0.65$).

Table 4

The correlations between speech-recognition scores and age at hearing-loss onset are generally smaller in magnitude than those listed in Table 3, suggesting that audibility and importance functions account for a larger proportion of the variance in the measured speech recognition-scores than age at hearing-loss onset, hearing-loss duration, or age. The difference between the mean (computed across the 500- to 4000-Hz CF range) of the Fisher-z-transformed correlation coefficients in Table 3 and the mean (computed across the same CF range) of the Fisher-z-transformed correlation coefficients for speech-recognition scores and age at hearing-loss onset in Table 4 was statistically significant ($p = 0.0264$; here, the uncorrected $p = 0.05$ level is used, since no multiple comparisons are involved).

The duration of HA use was significantly, but modestly, correlated to the duration of hearing loss ($r = 0.30$, $p = 0.0145$), and the p values in the last column of Table 4 indicate that it was not significantly correlated to speech-recognition scores for any of the CFs tested.

Finally, correlations between pure-tone thresholds and the following variables: age, age

at hearing-loss onset, and hearing-loss duration, were computed, separately, for each ear and each pure-tone frequency. None of these correlations was found to be statistically significant at the Bonferroni-corrected p level ($p = 0.05/34 = 0.0015$), although p values lower than the uncorrected p level ($p = 0.05$) were observed for correlations between thresholds measured at frequencies comprised between 1000 Hz and 1400 Hz and the following variables: age and age at hearing-loss onset. None of the correlations involving pure-tone thresholds and hearing-loss duration reach statistical significance (even at the uncorrected p level of 0.05).

DISCUSSION

Can pure-tone audiograms be used to infer the intelligibility of lowpass-filtered speech in HI individuals?

An automatic statistical-clustering algorithm was applied on pure-tone thresholds to separate 86 HI ears (43 HI subjects) into seven groups typified by different audiograms, which ranged from steeply sloping to quasi-flat (Figs. 2 and 3). The results obtained using this approach showed that, on average, different audiogram configurations tended to be associated with different PFs for the recognition of lowpass-filtered speech, i.e., different patterns of PC as a function of filter CF (Figs. 4 and 5). To some extent, these correspondences were as expected based on audibility considerations and importance functions for speech perception. In particular, groups in which thresholds were less elevated than those of other groups having, on average, higher speech-recognition scores than the latter. Also, when mean thresholds at 250 Hz were equated (Fig. 3), shallower audiogram slopes were usually associated with steeper PFs and higher speech-recognition scores for CFs above 500 Hz, compared to groups with steeper slopes (Fig. 5). However, importantly, the results also showed substantial variability in PFs across ears within

the same audiogram-type group, which led to substantial overlap in PC scores across the groups. Such across-ear variability in PFs for speech recognition, which cannot be accounted for simply by differences in pure-tone thresholds, obviously limits the extent to which an HI individual's ability to understand lowpass speech can be inferred, using solely the audiogram.

In addition, the results indicate that speech-recognition scores measured using broadband speech signals do not provide a reliable indication of how speech-recognition performance degrades, as the lowpass-filter CF is lowered. Ears with similar scores in the broadband condition ($CF = 22.05$ kHz) could exhibit markedly different PFs, with some PFs increasing continuously across the entire range of CFs tested in this study, while others reached a plateau at relatively low CFs; when speech-recognition scores saturated, the CF at which they did so did not appear to vary systematically with the score measured in the broadband condition.

These findings underline the importance of measuring—rather than trying to infer, based on the audiogram—speech-recognition performance with lowpass-filtered stimuli, in all clinical or research contexts in which knowing how well an HI individual can recognize speech using only low-frequency information is potentially consequential, e.g., when trying to determine which type of auditory prosthesis is best suited for an HI patient. It is worth noting that the criteria that are currently in use in various countries to determine candidacy to implantation with a long- or short-electrode cochlear implant are based primarily, if not exclusively, on pure-tone thresholds and on recognition scores measured using broadband speech.

How well can lowpass-speech recognition in HI listeners be predicted based on speech audibility and importance functions?

Speech-intelligibility predictions based on audibility and speech-importance functions were

found to be significantly correlated with the measured speech-recognition scores. Importantly, this was the case even when the correlations were computed separately for each CF. Such CF-specific correlations were significant for all CFs except the lowest two, and they were relatively high (with r 's ranging from 0.58 to 0.65). This outcome indicates that a substantial proportion (i.e., between 34% and 42%) of the variability in speech-recognition scores across the HI ears tested in this study could be explained in terms of audibility and importance functions. This is consistent with previous findings, which have shown that a primordial determinant of differences in speech-recognition performance among hearing-impaired (HI) individuals is the audibility of spectral components of the speech signal that are important for intelligibility (e.g., Dubno et al., 1989; Hogan & Turner, 1998; Rankovic, 1991;1998;2002). However, a substantial proportion (40% at least) of the variance in the measured speech-recognition scores remained unaccounted for by audibility and importance functions. This, too, is consistent with previous findings (e.g., Baer et al., 2002; Ching et al., 1998; Dubno et al., 1989; Hogan & Turner, 1998; Kamm et al., 1985; Pavlovic, 1984).

Overall, these results indicate that other factors than audibility influence the ability of HI individuals to use low-frequency information in speech.

Relationships between speech-recognition performance and other variables than pure-tones and audibility

Significant correlations were found between speech-recognition scores and other variables than predicted scores based on audibility. In particular, negative correlations were found between speech-recognition scores and age at hearing-loss onset, and positive correlations were found between speech-recognition scores and hearing-loss duration. Importantly, these correlations

cannot be explained as by-products of pre-existing correlations between speech-recognition scores and age because the correlation coefficients for speech-recognition scores and age were generally smaller than the correlation coefficients for speech-recognition scores and age at hearing-loss onset or hearing-loss duration, and they were not statistically significant at the Bonferroni-corrected p level. Moreover, the correlations are unlikely to be by-products of correlations between either hearing-loss onset age or hearing-loss duration and pure-tone thresholds: even if the uncorrected level ($p = 0.05$) is assumed, correlations between pure-tone thresholds and age at hearing-loss onset were only observed for a small subset of test frequencies between 1000 and 1400 Hz, whereas significant correlations between age at hearing-loss onset and speech-recognition scores were observed for a wider range of CFs, from 353 Hz to 4000 Hz. These findings confirm and extend, in a larger group of HI subjects, the recent findings of Seldran et al. (2011).

As suggested by Seldran et al. (2011), a possible explanation for positive correlations between speech-recognition performance and hearing-loss duration, is that protracted exposure to degraded auditory inputs progressively leads listeners with severe-to-profound high-frequency hearing loss to learn to extract more useful information from the low frequencies, compared to listeners who do not have to rely so heavily on these low-frequencies for speech understanding. This possibility was suggested by other authors (e.g., Rhodes, 1966; Vestergaard, 2003; Moore & Vinay, 2009), and it is consistent with another finding of the present study, to which we return in the next section, namely: the fact that HI listeners with steeply sloping audiograms actually had higher speech-recognition scores than the NH listeners.

In this context, a possible explanation for the finding of *negative* correlations between speech-recognition performance and age at hearing loss onset is that individuals who have been

hearing undistorted speech signals for several years before they incurred a hearing loss may need more time to “unlearn” implicit speech-recognition strategies that rely heavily on medium or high frequencies, where spectral components are now less audible, and distorted. Further work is needed to determine whether this explanation is correct. In particular, studies measuring frequency-weighting functions for speech recognition using a correlational method (Doherty & Turner, 1996; Turner et al., 1998) in HI listeners, may provide further insight into this issue.

Other possible explanations for the existence of significant correlations between speech-recognition performance and hearing loss onset age, or hearing-loss duration, were mentioned by Seldran et al. (2011). One of these explanations relates to differences in hearing-loss etiology depending on the stage of life at which the hearing loss is incurred. For example, hearing-loss acquired at a late age is more likely to be due to noise exposure or ischemia than congenital hearing loss, and hearing loss incurred during childhood is more likely to be related to ear infections (otitis media) or measles, than is hearing loss incurred in adulthood. To the extent that the physiological mechanisms responsible for hearing loss associated with different etiologies affect the ability to understand speech in a different way or to a different degree, these differences could play a role in the observed correlations. The etiology of the hearing loss for the ears tested in this study, when available, is reported in Table I. Unfortunately, the information was too scant to allow us to explore the role of differences in etiology as a possible factor of speech-recognition performance.

Better speech recognition in ears with steeply-sloping hearing-loss than in NH ears?

A marginal but interesting outcome of this study relates to the observation that, in ears with good residual hearing at low frequencies and relatively steep audiogram slopes, the mean speech-

recognition scores for CFs of 500 and 707 Hz were often higher than the highest mean recognition score measured at the corresponding CF in the NH listeners (compare the blue and green curves in the two upper panels in Fig. 4 to the highest dashed line). Rhodes (1966) and Vestergaard (2003) also found that some HI listeners with high-frequency hearing loss had higher speech-recognition scores for lowpass-filtered speech than NH listeners, or, than expected based on audibility. Moore & Vinay (2009) found that listeners who had cochlear “dead regions” at high frequencies were better able to recognize lowpass-filtered speech signals than listeners who did not have dead regions at higher frequencies. Vestergaard (2003) suggested that listeners who have cochlear dead regions are accustomed to relying exclusively on low-frequency acoustic cues for speech perception. Although the listeners in the present study were not tested for cochlear dead regions, considering that dead regions are almost always present at frequencies where pure-tone thresholds are larger than 80-90 dB HL (see: Moore, 2001), it seems highly likely that most of the ears in the “steeply sloping” and “moderately sloping” groups of the current study almost certainly had dead regions at high frequencies.

Regardless of whether or not dead regions were present at high frequencies in these ears, the finding of higher speech-recognition scores in these HI ears than in NH ears cannot be accounted for entirely by differences in speech audibility between the two groups. Although the HI ears were tested using higher speech-presentation levels than the highest level used in NH listeners, the latter level (80 dB SPL) was sufficiently high to ensure that the low-frequency components of speech (below 500 Hz) were well audible for the NH listeners. Therefore, another explanation is needed, for why HI subjects with severe or profound hearing loss at high frequency but relatively good residual hearing at lower frequencies are better able to recognize lowpass-filtered speech than NH subjects. A possible explanation for this outcome is that these

HI subjects have more extensive experience relying primarily or exclusively on low-frequency cues for speech understanding than NH listeners. This explanation is consistent with the explanation suggested in the previous section, for why speech recognition scores are negatively correlated with hearing-loss duration.

ACKNOWLEDGEMENTS

This work was supported by the French National Center for Scientific Research (CNRS), Vibrant Med-El France (Doctoral Research Grant CIFRE 266/2007 to FS), and an NIH grant R01 DC 05216 (CM). Subsets of the data presented in this manuscript were used in two previously published studies (Seldran et al., 2011; Seldran et al., 2010). The authors would like to thank Michel Belaieff, Christian Berger-Vachon, Lionel Collet, Samia Labassi and Bénédicte Philibert for their support and advice.

References

- Allen PD, Eddins DA (2010) Presbycusis phenotypes form a heterogeneous continuum when ordered by degree and configuration of hearing loss. *Hear Res* 264:10–20
- ANSI. ANSI S3.5 (1969) Methods for the Calculation of the Speech Articulation Index. New York: ANSI
- ANSI. ANSI S3.5 (1997) Methods for the Calculation of the Speech Intelligibility Index. New York: ANSI
- Baer T, Moore BC, Kluk K (2002) Effects of low pass filtering on the intelligibility of speech in noise for people with and without dead regions at high frequencies. *J Acoust Soc Am* 112:1133-44
- Ching TY, Dillon H, Byrne D (1998) Speech recognition of hearing-impaired listeners: predictions from audibility and the limited role of high-frequency amplification. *J Acoust Soc Am* 103:1128-1140
- Doherty KA, Turner CW (1996) Use of a correlational method to estimate a listener's weighting function for speech. *J Acoust Soc Am* 100:3769-73
- Dubno JR, Dirks DD, Schaefer AB (1989) Stop-consonant recognition for normal-hearing listeners and listeners with high-frequency hearing loss. II: Articulation index predictions. *J Acoust Soc Am* 85:355-64
- Duda RO, Hart PE, Stork DG (2001) Pattern classification (2nd ed.) Wiley, New York, NY
- Efron B, Tibshirani R (1993). An Introduction to the Bootstrap. Chapman and Hall/CRC, New York, NY
- Fisher RA (1915) Frequency distribution of the values of the correlation coefficient in samples of an indefinitely large population. *Biometrika* 10:507-521

- Fournier JE (1951) *Audiométrie vocale* Maloine, Paris
- Gantz BJ, Hansen MR, Turner CW, Oleson JJ, Reiss LA, Parkinson AJ (2009) Hybrid 10 clinical trial: preliminary results. *Audiol Neurotol* 14 Suppl 1:32-8
- Hogan CA, Turner CW (1998) High-frequency audibility: benefits for hearing-impaired listeners. *J Acoust Soc Am* 104:432-41
- Kamm CA, Dirks DD, Bell TS (1985) Speech recognition and the Articulation Index for normal and hearing-impaired listeners. *J Acoust Soc Am* 77:281-8
- Lloyd SP (1982) Least squares quantization in PCM. *IEEE Transactions on Information Theory* 28:129–137
- Moore BC, Vinay SN (2009) Enhanced discrimination of low-frequency sounds for subjects with high-frequency dead regions. *Brain* 132:524-536
- Moore BCJ (2001) Dead regions in the cochlea: Diagnosis, perceptual consequences, and implications for the fitting of hearing aids. *Trends Amplif* 5:1-34
- Murray N, Byrne D (1986) Performance of hearing-impaired and normal hearing listeners with various high-frequency cut-offs in hearing aids. *Austr J Audiol* 8:21-28
- Pavlovic CV (1984) Use of the articulation index for assessing residual auditory function in listeners with sensorineural hearing impairment. *J Acoust Soc Am* 75:1253-1258
- Rankovic CM (1991) An application of the articulation index to hearing aid fitting. *J Speech Hear Res* 34:391-402
- Rankovic CM (1998) Factors governing speech reception benefits of adaptive linear filtering for listeners with sensorineural hearing loss. *J Acoust Soc Am* 103:1043-1057
- Rankovic CM (2002) Articulation index predictions for hearing-impaired listeners with and without cochlear dead regions. *J Acoust Soc Am* 111:2545-2548

- Rhodes RC (1966) Discrimination of filtered CNC lists by normals and hypacusics. *J Aud Res* 6:129-133
- Seldran F, Gallego S, Micheyl C, Veuillet E, Truy E, Thai-Van H (2011) Relationship Between Age of Hearing-Loss Onset, Hearing-Loss Duration, and Speech Recognition in Individuals with Severe-to-Profound High-Frequency Hearing Loss. *J Assoc Res Otolaryngol* 12:519-34
- Seldran F, Thai-Van H, Truy E, Berger-Vachon C, Collet L, Gallego S, Beliaeff M (2010) A Filtered Speech Test to Better Evaluate Electric Acoustic Stimulation (EAS) Candidacy. *Cochlear Implants Int* 11 Suppl 1:130-3
- Skinner MW, Miller JD (1983) Amplification bandwidth and intelligibility of speech in quiet and noise for listeners with sensorineural hearing loss. *Audiology* 22:253-79
- Steinhaus H (1956) Sur la division des corps matériels en parties. *Bull Acad Polon Sci* 4:801-804
- Studebaker GA (1985) A "rationalized" arcsine transform. *J Speech Hear Res* 28:455-62
- Turner CW (2006) Hearing loss and the limits of amplification. *Audiol Neurotol* 11 Suppl 1:2-5
- Turner CW, Cummings KJ (1999) Speech audibility for listeners with high-frequency hearing loss. *Am J Audiol* 8:47-56
- Turner CW, Kwon BJ, Tanaka C, Knapp J, Hubbartt JL, Doherty KA (1998) Frequency-weighting functions for broadband speech as estimated by a correlational method. *J Acoust Soc Am* 104:1580-5
- Vestergaard MD (2003) Dead regions in the cochlea: implications for speech recognition and applicability of articulation index theory. *Int J Audiol* 42:249-61

- Vickers DA, Moore BC, Baer T (2001) Effects of low-pass filtering on the intelligibility of speech in quiet for people with and without dead regions at high frequencies. *J Acoust Soc Am* 110:1164-75
- Villchur E (1973) Signal processing to improve speech intelligibility in perceptive deafness. *J Acoust Soc Am* 53:1646-57
- Woodson EA, Reiss LA, Turner CW, Gfeller K, Gantz BJ (2010) The Hybrid cochlear implant: a review. *Adv Otorhinolaryngol* 67:125-34

Figure captions

Figure 1. Individual and mean audiograms of the 43 HI listeners. Data for the left and right ears are shown separately in the left and right panels, respectively. The gray lines show individual data. The solid black lines shows the mean across listeners.

Figure 2. Mean and individual audiograms for the seven groups of ears separated by the k-means algorithm. The panels on the left show group means, and 95% confidence intervals around the group mean, separately for each pure-tone frequency. The panels on the right show “individual” data, i.e., one audiogram per ear. The top row shows the three groups with the steepest audiogram slopes. The middle row shows three groups with moderate slopes. Within each of these two rows, the different groups are indicated by different colors: green for the steeper-slope group; red for the shallower-slope group; blue for the intermediate-slope group. The bottom row shows the group of ears with quasi-flat audiograms. Note that the same color (e.g., blue) is used to indicate different groups across different panels.

Figure 3. Mean and individual audiograms of two pairs of groups with similar mean thresholds at 250 Hz, but different slopes. These data are replotted from Fig. 2, to facilitate visual comparisons of slopes between groups with similar mean thresholds at low frequencies, but different slopes.

Figure 4. Mean and individual PFs, showing PC as a function of lowpass-filter CF, for the seven groups of ears formed on the basis of pure-tone thresholds. The ordering of the panels, the colors, and the symbols are the same as in Fig. 2. The error bars show 95% confidence intervals around the group mean. The dashed lines show mean PFs across 20 NH listeners, replotted from Seldran et al. (2011). These “reference” PFs were measured using the same speech stimuli, and presentation levels ranging from 10 to 80 dB SPL in 10-dB steps. However, since for levels of 40

dB SPL and higher, the PFs were very similar and overlapped largely, they were averaged; this is why only four dashed curves are shown here, corresponding to the PFs for 10, 20, and 30 dB SPL, and to the average PF across 40-80 dB SPL, from the lowest to the highest.

Figure 5. Mean and individuals PFs for the two pairs of groups whose audiograms are shown in Fig. 3. These data are replotted from Fig. 4 to facilitate comparisons of the PFs across groups with similar mean pure-tone thresholds at low frequencies but different audiogram slopes.

Figure 6. Scatter plot of the predicted versus measured rau scores. The different combinations of symbols and colors correspond to different CF conditions, as indicated in the key. The diagonal line shows where data points would fall if the predicted scores were equal to the measured scores.

Table 1. Participant information. In the first column, participants are numbered based on age at testing, from youngest to oldest. Stars are used to indicate participants whose data were used in a previous study (Seldran et al., 2011). The participant’s gender (M: male; F: female) and age in years (yr.) are indicated in the second and third columns. The column labeled “Side” indicates whether the information listed in the following columns is for the right (R) or the left (L) ear. The column labeled “HL dur. (yr.)” indicates the duration of hearing loss. The column labeled “HA dur. (yr.)” indicates the number of years of regular HA use. The column labeled “Speech level” indicates the level at which the speech stimuli were presented, in the considered ear. The columns labeled “PTA” list pure-tone threshold averages across the indicated frequency range. In “Etiology” column, the abbreviations “Pott’s” and “Usher” stand for “Pott’s disease” and “Usher’s syndrome,” respectively.

Sbj.	Gender	Age (yr.)	Side	HL dur. (yr.)	HA dur. (yr.)	Speech level (dBSPL)	PTA 250-850 Hz (dB HL)	PTA 1-4 kHz (dB HL)	Etiology
1	M	21	R	21	16	90	71	104	congenital
			L	21	16	95	74	107	congenital
2	M	22	R	22	17	100	79	101	congenital
			L	22	17	100	79	108	congenital
3*	M	29	R	28	18	100	81	105	congenital
			L	28	18	100	78	101	congenital
4	F	30	R	5	5	100	71	94	unknown
			L	5	5	100	101	106	unknown
5*	F	30	R	15	0	90	67	111	unknown

			L	15	5	90	69	118	unknown
6*	M	31	R	31	21	80	33	105	unknown
			L	31	21	80	36	107	unknown
7	F	32	R	32	29	100	79	110	congenital
			L	32	29	100	93	116	congenital
8	F	33	R	33	4	100	84	109	congenital
			L	33	4	100	100	122	congenital
9	F	34	R	34	4	90	14	106	congenital
			L	34	4	90	38	112	congenital
10	F	36	R	36	0	100	63	110	rubella
			L	36	0	95	68	115	rubella
11*	F	37	R	6	0	80	53	111	unknown
			L	6	0	100	75	124	unknown
12	F	38	R	32	32	100	72	112	unknown
			L	32	32	100	69	112	unknown
13	F	38	R	38	2	90	54	95	congenital
			L	38	2	80	49	96	congenital
14*	F	38	R	38	11	100	85	117	Usher
			L	38	11	100	69	112	Usher
15	M	47	R	5	4	90	34	113	ototoxicity
			L	5	0	100	89	124	ototoxicity
16	F	48	R	48	42	100	84	86	congenital
			L	48	7	100	75	84	congenital

17	F	51	R	21	12	80	30	109	unknown
			L	21	12	90	31	104	unknown
18*	F	53	R	28	18	100	78	100	unknown
			L	28	18	100	84	102	unknown
19	M	54	R	9	0	100	100	113	trauma
			L	8	8	100	73	110	unknown
20	M	54	R	11	7	100	74	83	unknown
			L	5	5	100	76	85	unknown
21*	F	54	R	50	45	100	71	110	congenital
			L	50	45	100	71	108	congenital
22	F	56	R	13	13	100	61	109	ototoxicity
			L	13	13	100	53	110	ototoxicity
23	F	56	R	31	15	100	71	76	unknown
			L	31	15	100	73	99	unknown
24*	F	59	R	4	4	100	66	120	unknown
			L	4	4	90	77	123	unknown
25	F	59	R	19	13	100	54	117	unknown
			L	19	13	80	55	102	unknown
26	F	59	R	25	20	80	49	104	unknown
			L	25	20	90	85	123	unknown
27	M	59	R	49	9	90	61	102	unknown
			L	49	9	100	61	106	unknown
28*	M	60	R	22	12	100	76	119	unknown

			L	22	12	95	73	116	unknown
29*	M	63	R	7	7	100	91	125	head injury
			L	44	3	100	44	114	head injury
30*	F	64	R	64	10	80	29	112	Pott's
			L	64	0	80	41	122	Pott's
31	M	65	R	25	17	100	62	124	unknown
			L	25	17	100	59	122	unknown
32	F	65	R	61	19	100	78	122	unknown
			L	61	0	100	86	117	unknown
33*	M	68	R	18	14	100	74	98	unknown
			L	18	14	100	78	94	unknown
34*	M	69	R	9	3	100	56	93	unknown
			L	9	3	100	68	102	unknown
35	M	69	R	20	0	100	120	125	trauma
			L	1	0	100	64	78	unknown
36	F	71	R	26	26	95	83	87	unknown
			L	26	26	100	85	122	unknown
37*	M	71	R	41	18	100	54	111	unknown
			L	41	18	100	64	118	unknown
38	F	72	R	16	16	100	86	112	unknown
			L	16	16	100	83	122	unknown
39	F	72	R	44	0	90	78	72	trauma
			L	1	1	90	94	69	unknown

40	F	73	R	19	4	100	69	117	ototoxicity
			L	19	0	100	106	124	ototoxicity
41	F	73	R	25	18	90	61	114	unknown
			L	25	18	90	59	112	unknown
42	F	74	R	14	8	100	48	93	unknown
			L	14	3	90	39	104	unknown
43	M	74	R	44	7	90	63	66	unknown
			L	44	7	90	66	64	unknown

Table 2. Audiogram slopes for the seven groups of ears. In the first column, the groups are ordered by slope, with group 1 corresponding to the steepest slope, and group 7 corresponding to the shallowest (quasi-flat) slope. The second column indicates the number of subjects in each group. The third column indicates the type of audiogram slope for the considered group: steep, moderate, and quasi-flat; these three types of slopes correspond to the three rows in Fig. 2. The fourth column indicates the color that was used to represent the audiogram of the considered group in Figs. 2 to 5. The fifth column indicates the mean slope (in dB/octave) measured over the 250-1400-Hz range, i.e., a 2.5-octave range, in the considered group; the frequency of 1400 Hz was chosen as the upper limit for the range because it usually coincided with a knee-point in the group-mean pure-tone threshold curves (left-hand panels in Fig. 2). The sixth column lists the group-mean pure-tone thresholds at 250 and 1400 Hz, and the difference between these two thresholds, which corresponds to the average “drop” in threshold between 250 and 1400 Hz. The last column lists the group-mean slope (in dB/octave) over the 1.4-4 kHz range.

Group	N	Slope type	Color	0.25 - 1.4 kHz		1.4 - 4 kHz
				Slope (dB/oct.)	Drop (dB)	Slope (dB/oct.)
1	10	steep	green	39.2	12 – 111 = 98	3.7
2	19		blue	30.5	38 – 114 = 76	3.7
3	11		red	25.1	60 – 122 = 62	0.9
4	9	moderate	green	24.0	39 – 96 = 60	11.4
5	20		blue	16.3	62 – 103 = 41	6.3
6	6		red	9.7	90 – 114 = 24	2.8
7	11	quasi flat	blue	3.5	67 – 76 = 9	4.6

Table 3. Pearson’s correlation coefficients (r) between measured and predicted rau scores for each CF. The first column, labeled “CF”, indicates the lowpass-filter cutoff frequency. Significant correlations at the Bonferroni-corrected p-level ($p = 0.05/8 = 0.00625$) are shown in bold. The p values listed in this table were computed separately for each set of variables, using a permutation procedure (see Data Analysis); as a result, it is possible for the same r value to be associated with different p values.

CF (Hz)	Predicted score	
	r	p
250	-0.35	0.0009
353	-0.18	0.0940
500	0.60	p<0.0001
707	0.63	p<0.0001
1000	0.65	p<0.0001
1414	0.62	p<0.0001
2000	0.58	p<0.0001
22050	0.59	p<0.0001

Table 4. Correlation coefficients between rau scores and age, age at hearing-loss onset, hearing-loss duration, and duration of HA use. Significant correlations at the Bonferroni-corrected p -level ($p = 0.05/8 = 0.00625$) are shown in bold. The p values listed in this table were computed separately for each set of variables, using a permutation procedure (see Data Analysis); as a result, it is possible for the same r value to be associated with different p values.

CF (Hz)	Age		HL age		HL duration		HA duration	
	r	p	r	p	r	p	r	p
250	-0.07	0.2887	-0.25	0.0234	0.29	0.0179	0.08	0.2811
353	-0.26	0.0273	-0.35	0.0038	0.23	0.0529	0.13	0.2117
500	-0.18	0.0962	-0.41	0.0006	0.39	0.0014	0.09	0.2925
707	-0.23	0.0410	-0.46	0.0001	0.41	0.0003	0.15	0.1289
1000	-0.25	0.0230	-0.40	0.0004	0.31	0.0047	0.20	0.0483
1414	-0.24	0.0322	-0.47	0.0001	0.41	0.0003	0.23	0.0325
2000	-0.21	0.0488	-0.36	0.0011	0.30	0.0050	0.19	0.0608
22050	-0.19	0.0795	-0.35	0.0017	0.31	0.0069	0.19	0.0686

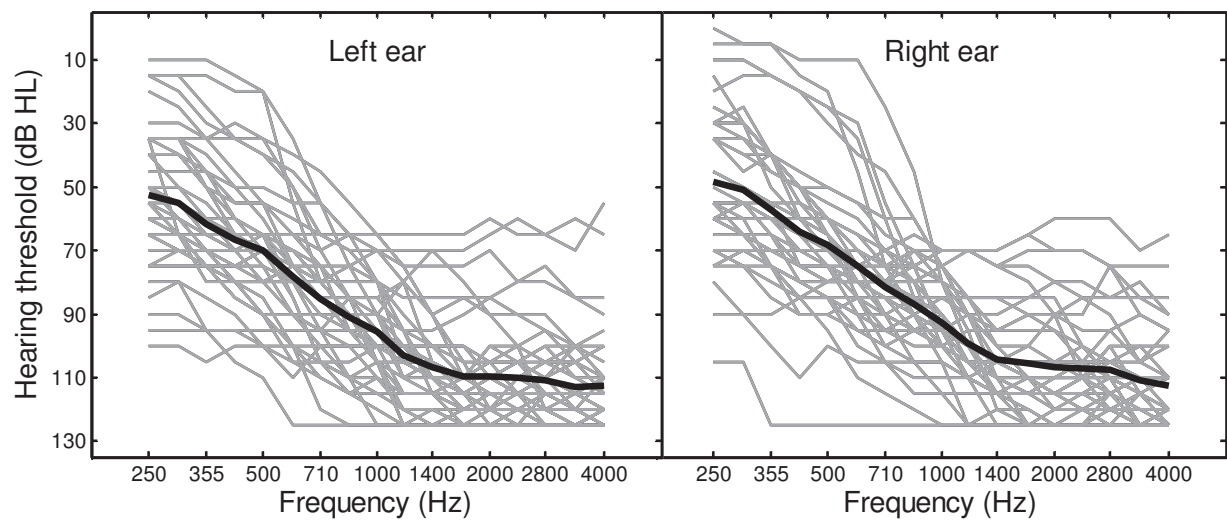


Figure 1

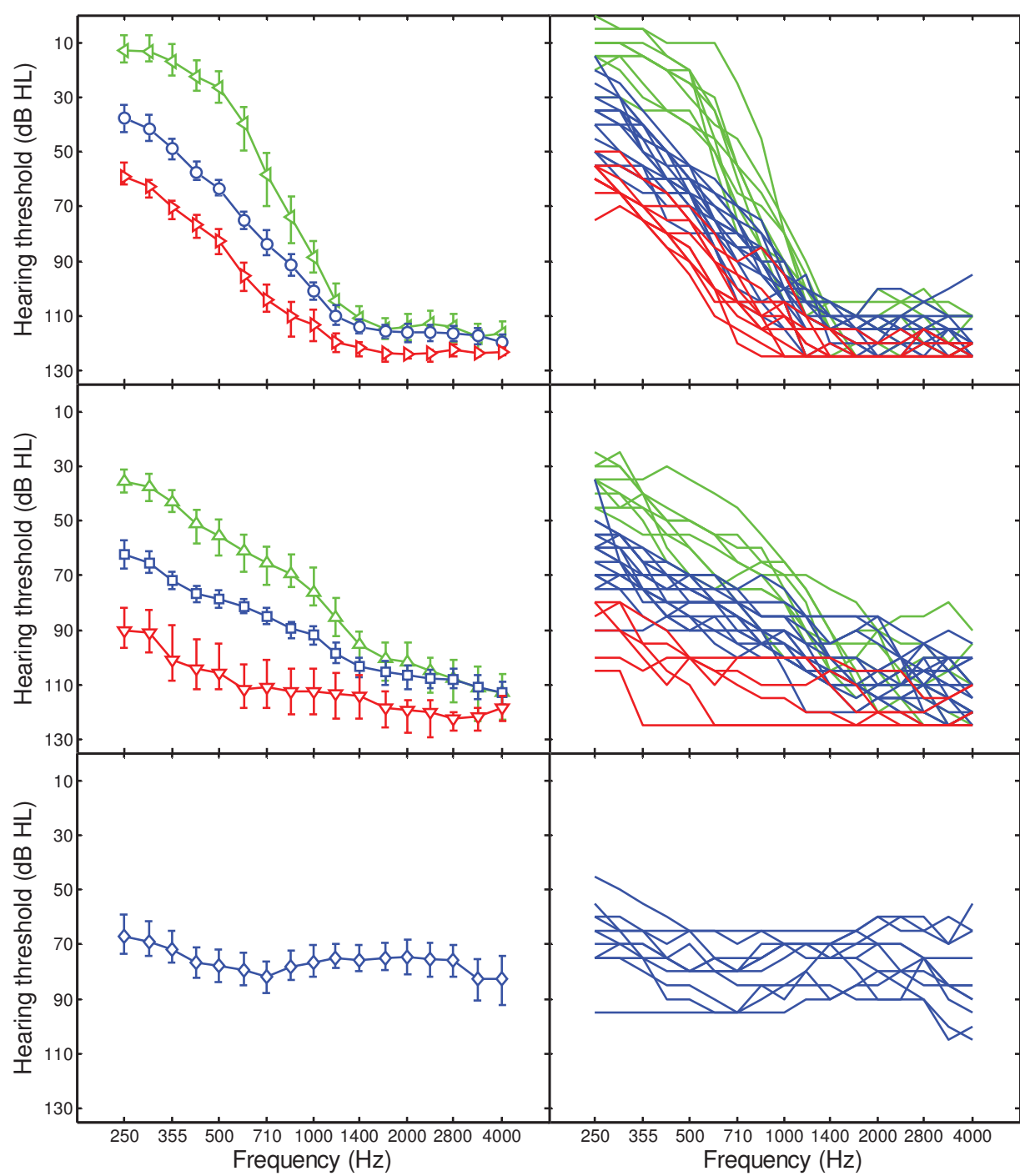


Figure 2

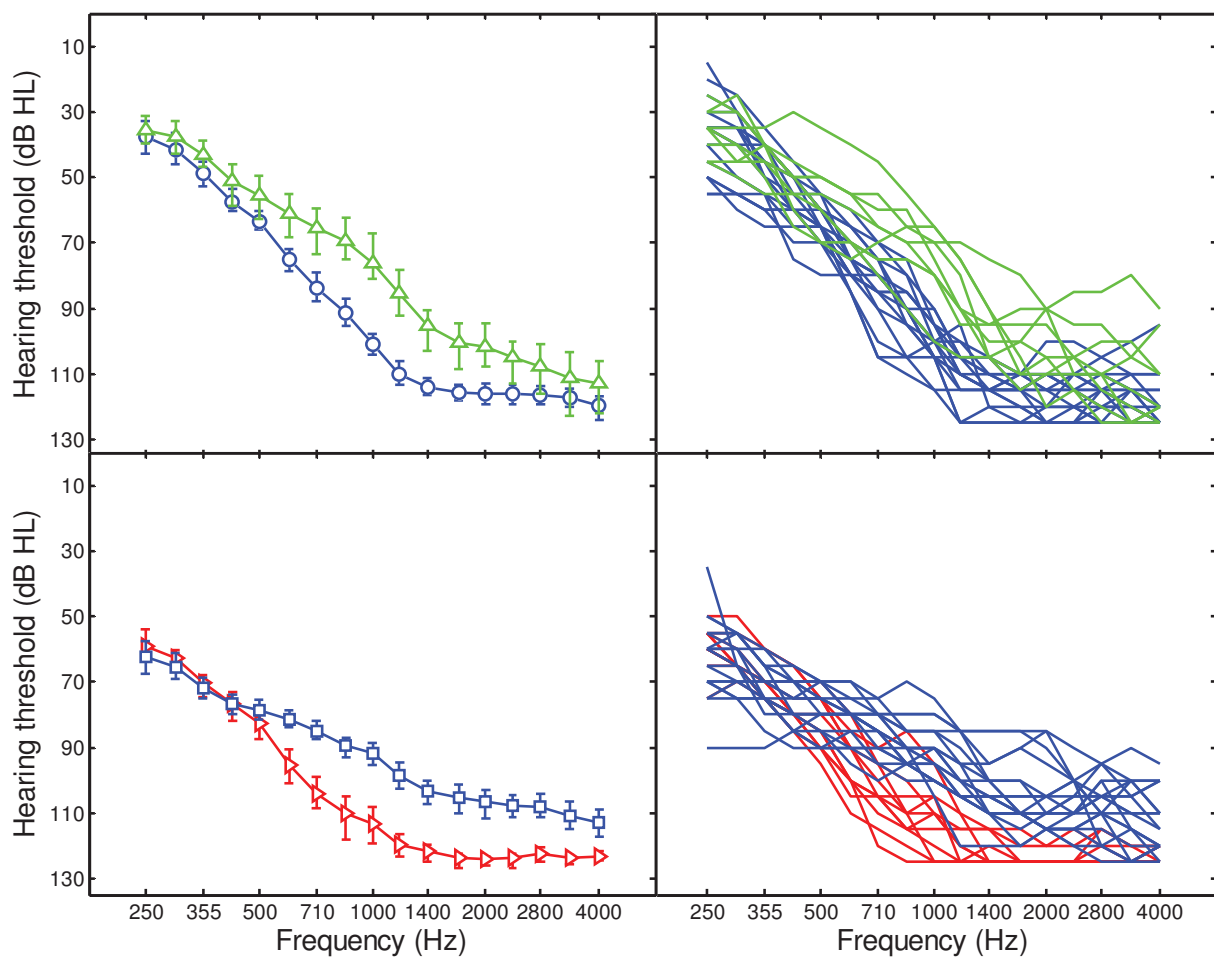


Figure 3

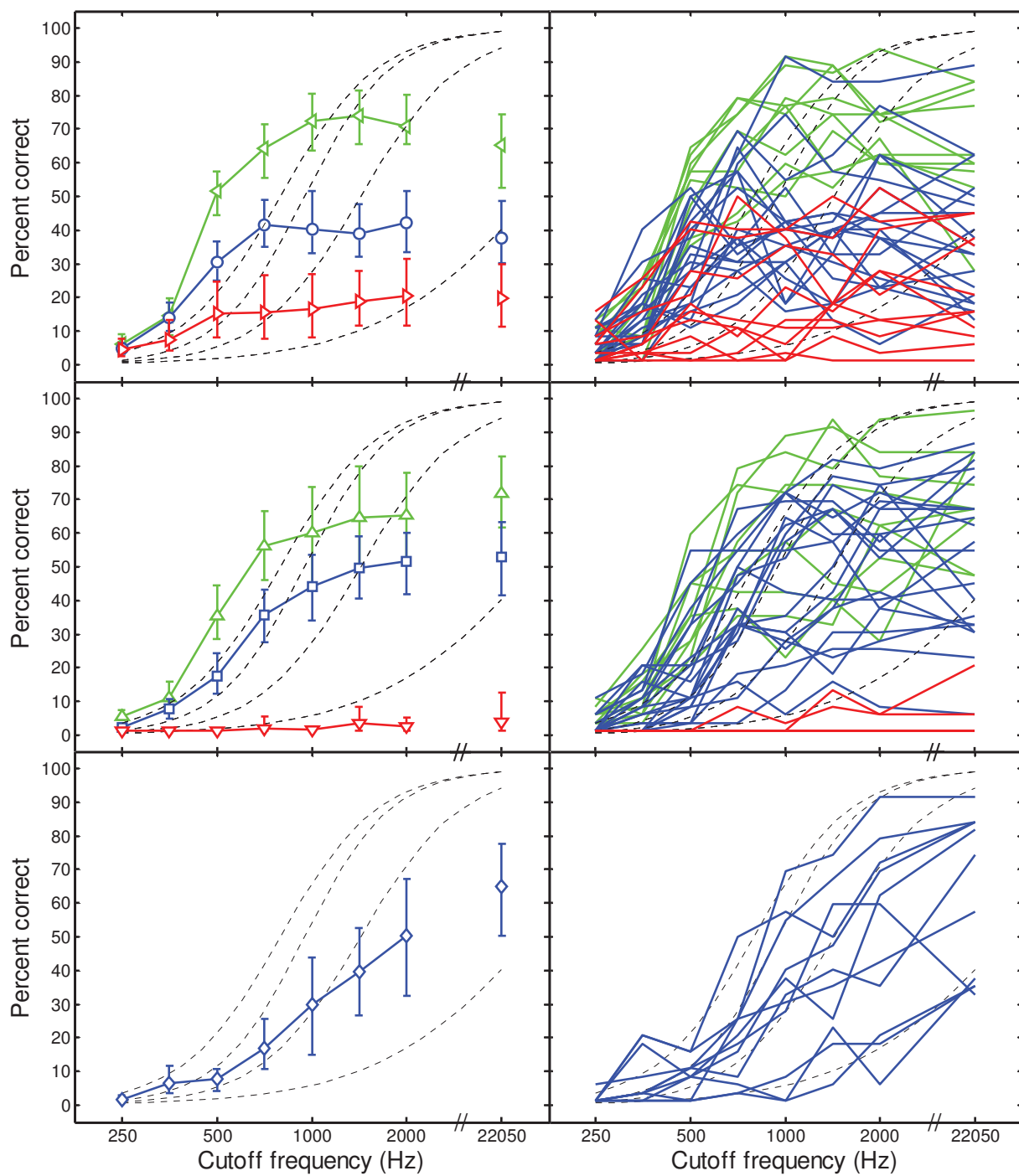


Figure 4

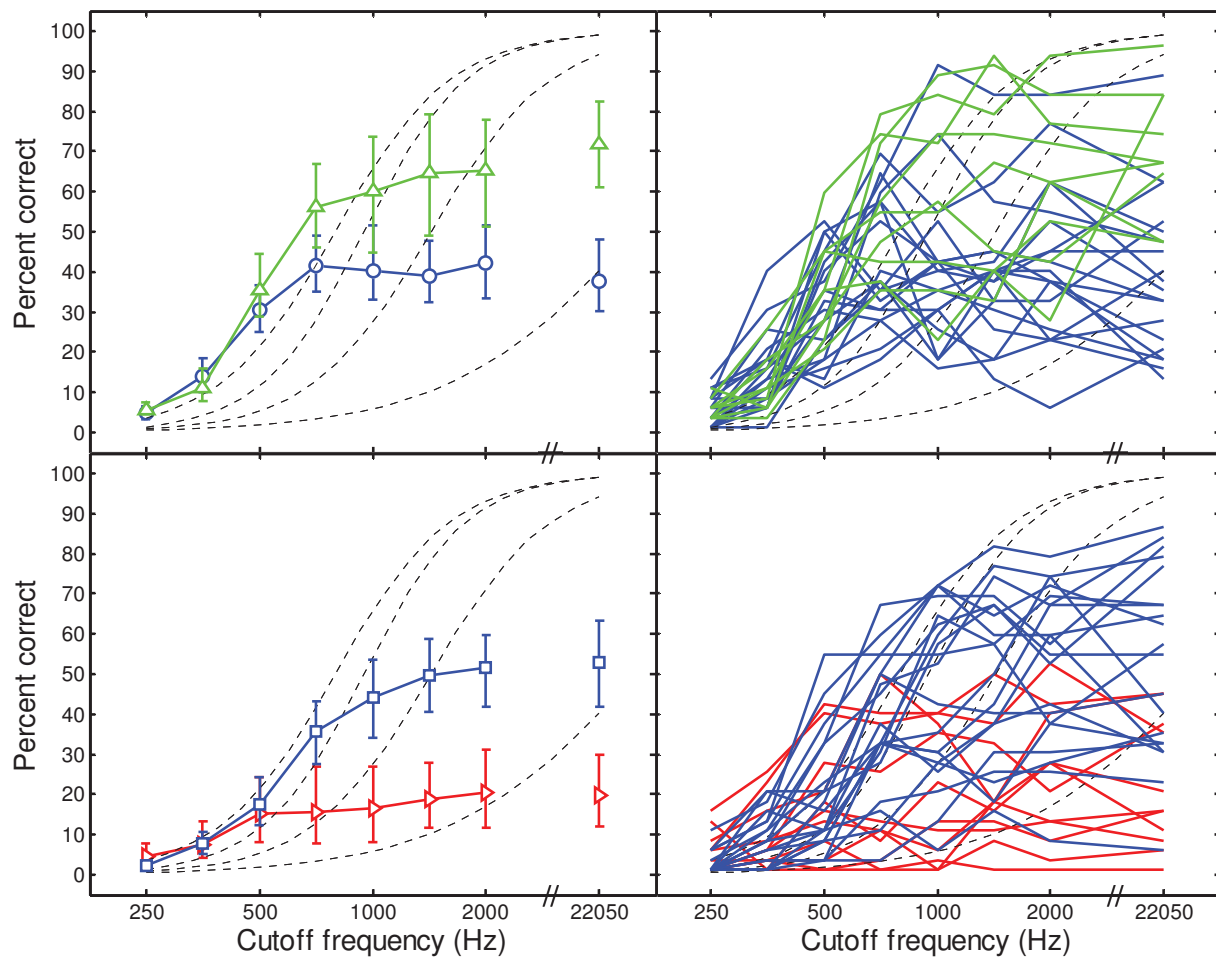


Figure 5

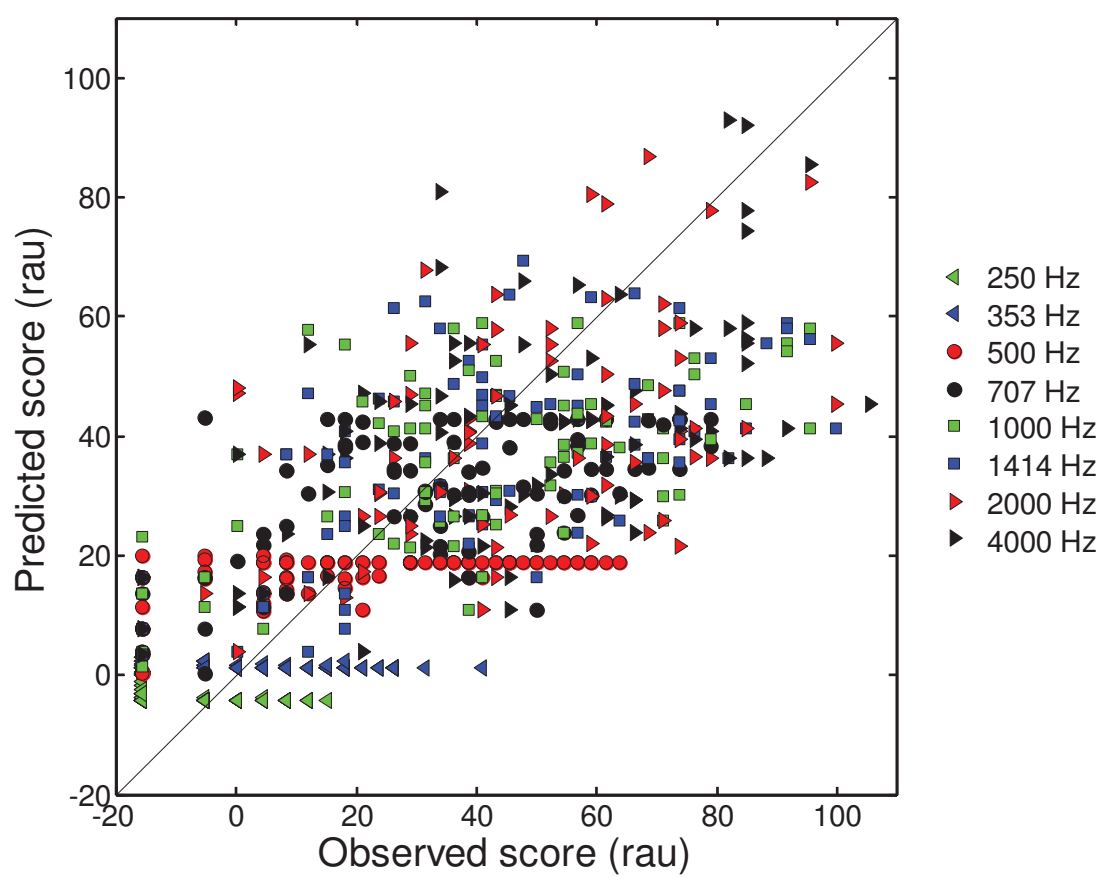


Figure 6

II.4. Suivi longitudinal de 8 patients EAS : expérience de Lyon

Dans cette section nous allons présenter les données pré- et post-implantation recueillies chez 8 sujets que nous avons suivis à Lyon sur une durée de 2 ans. Les 8 sujets étaient potentiellement candidats à la stimulation bimodale d'après leur audiométrie tonale et vocale et ont été implantés dans le cadre d'un protocole EAS avec des implants Med-El Sonata ti100 (12 canaux de stimulation). Les sujets ont bénéficié d'une chirurgie mini-invasive, en vue d'une préservation de leur audition. Sept sujets ont reçu une électrode Flex EAS (24 mm) et un sujet (la 2^{ème} patiente) a reçu une électrode Flex Soft (31 mm). Six sujets ont eu une insertion partielle du porte-électrode avec 2 électrodes extra-cochléaires, ce qui correspond à une insertion de 18-20 mm soit environ la longueur du tour basal. Les deux sujets restants ont eu une insertion totale.

Dans un premier temps, nous allons montrer l'évolution, au cours des 2 premières années d'utilisation, des seuils de perception et d'inconfort à l'audiométrie tonale, oreilles nues. Nous présenterons ensuite l'évolution des impédances des électrodes et des seuils maxima de confort mesurés chez les sujets. Dans un deuxième temps nous présenterons les données individuelles pour les résultats à l'audiométrie tonale et vocale.

La figure 12 ci-dessous montre l'évolution des seuils de perception (pré-implantation, à J0 ou J7, à J 1mois, à J 3mois, à J 6mois, à J 1an et à J 2ans) et d'inconfort (pré-implantation et à J 2ans) chez nos 8 sujets implantés. Les seuils de perception semblent se dégrader immédiatement d'environ 20 dB après la chirurgie et restent assez stables au cours des 2 premières années. Le seuil d'inconfort quant à lui s'est élevé d'une dizaine de dB entre les mesures pré-implantation et J 2ans. En terme de dynamique, les sujets n'ont donc perdu qu'une dizaine de dB avec l'implantation. L'élévation du seuil d'inconfort s'explique probablement par un rinne dû à la présence du porte-électrode dans la rampe tympanique.

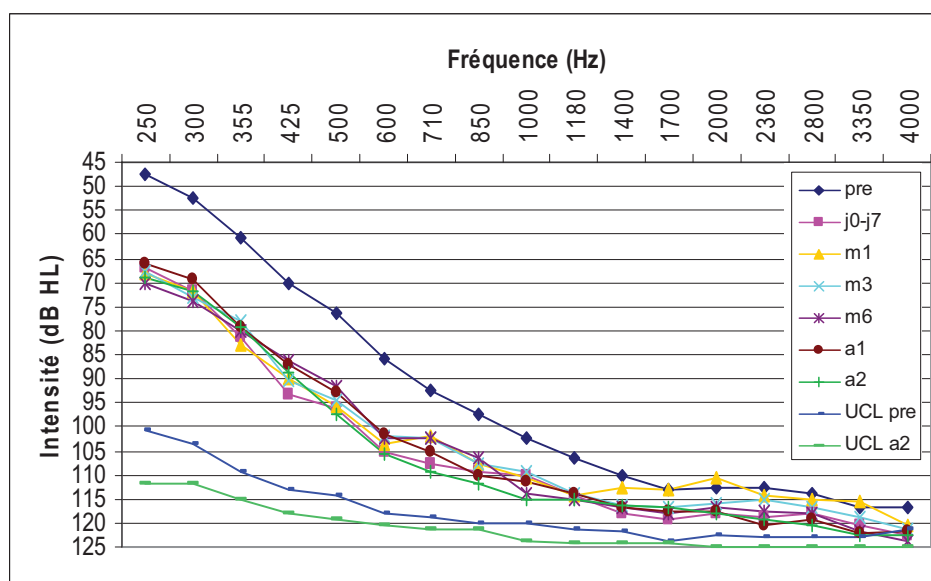


Figure 12. Evolution des seuils de perception et d'inconfort (oreilles nues au casque) au cours des 2 premières années d'implantation (n = 8).

La figure 13 représente l'évolution des impédances (exprimées en $k\Omega$) au cours des deux premières années d'utilisation de l'implant. Nous constatons que les impédances diminuent au cours du temps pour les électrodes E1 à E8. Au contraire, les impédances semblent augmenter pour les électrodes basales E9 et E10. Cette augmentation des impédances des électrodes basales s'explique sans doute par la formation de nouvel os et de tissu fibreux près du site d'insertion de l'électrode (Lie et al., 2007 ; Somdas et al., 2007).

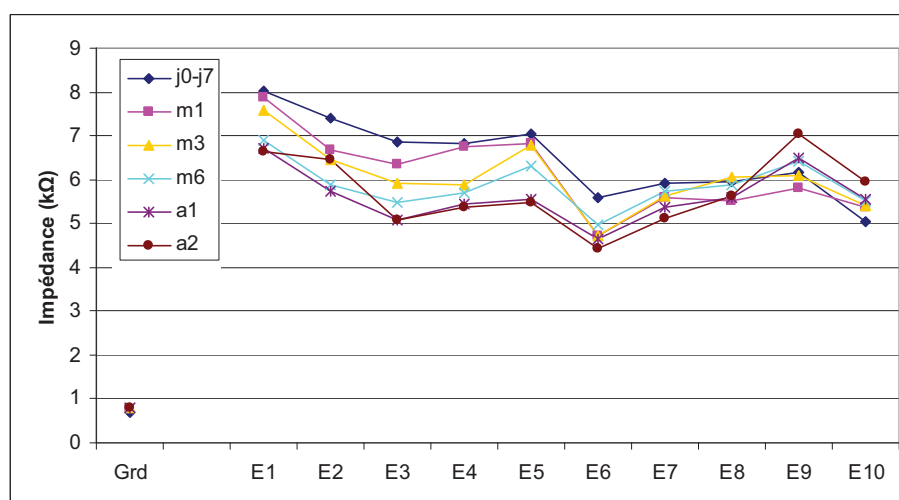


Figure 13. Evolution des impédances (exprimées en $k\Omega$) au cours des 2 premières années d'implantation (n = 8).

La figure 14 représente l'évolution des seuils maxima de confort (exprimés en nC) au cours des deux premières années d'utilisation de l'implant. L'augmentation des seuils est assez régulière.

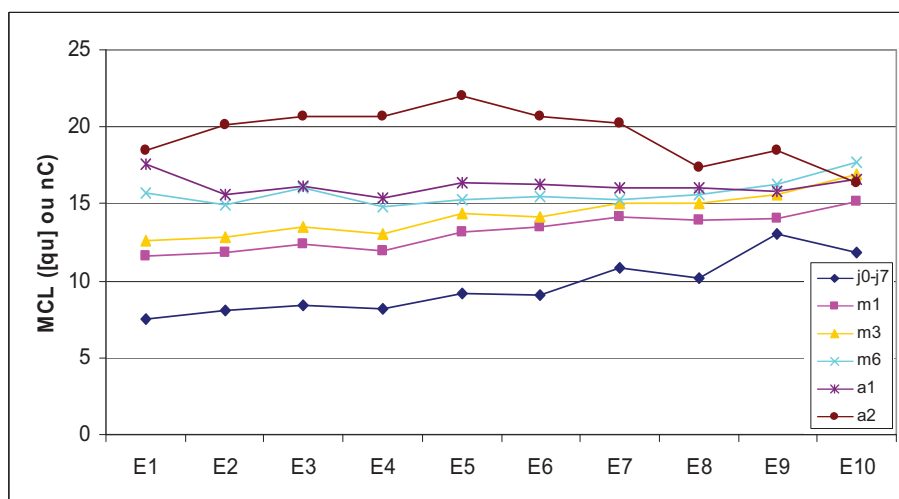


Figure 14. Evolution des seuils maxima de confort MCL (exprimés en unités [qu] équivalentes à des nC) au cours des 2 premières années d'implantation (n = 8).

Nous allons maintenant décrire les résultats individuels de nos résultats implantés EAS. Les patients sont présentés dans l'ordre chronologique où ils ont été implantés. L'information sur la durée de privation est en réalité la durée de surdité sévère/profonde dans les hautes fréquences. Certains des sujets ont ainsi une surdité qui est apparue dans la petite enfance mais c'est probablement leur audition résiduelle dans les basses fréquences qui leur a permis d'acquérir le langage.

Pour chaque patient, nous montrons 3 ou 4 graphes. Sur les figures suivantes nous trouverons en haut à gauche l'audiométrie tonale des patients avec les audiogrammes pré-implantation, post-implantation (seuil de perception et seuil d'inconfort), seuil de perception avec la stimulation acoustique seule et en condition EAS.

Les cadres en haut à droite montrent les résultats à l'audiométrie vocale filtrée passe-bas obtenus en pré-implantation et post-implantation lorsque cela était possible.

Les cadres en bas à gauche montrent les scores obtenus à l'audiométrie vocale entre J7 et J 2ans, pour les conditions EAS, EAS + aide auditive controlatérale (EAS + HA), acoustique seul (A only), électrique seul (E only) et aide auditive controlatérale seule (HA oppos only).

Enfin les cadres en bas à droite présentent les résultats obtenus pour les tests de résistivité au bruit speech noise, dans les meilleures conditions de test (i.e. implant EAS + prothèse controlatérale si le patient en a une). Pour évaluer où se situent nos implantés pour la compréhension dans le bruit par rapport à l'implanté cochléaire, nous nous sommes basés sur un article de Turner et al. (2008) qui trouvent des seuils de compréhension de la parole autour de 8 dB (figure 15). Dans l'article 2 du présent manuscrit où nous avons testé une cinquantaine de patients implantés, nous montrons que les utilisateurs d'un implant cochléaire unilatéral ont un seuil à 50% obtenu à un RSB = 5,5 dB ; pour les utilisateurs d'un implant cochléaire et d'une prothèse controlatérale, le seuil à 50 % est obtenu pour un RSB = 3,1 dB.

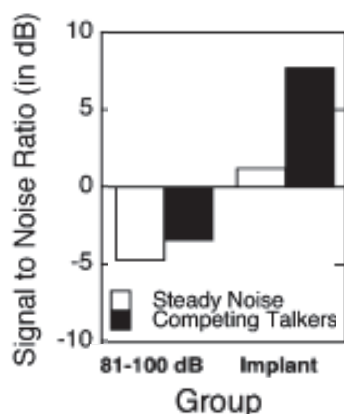


Figure 15. Seuils de reconnaissance de la parole dans le bruit continu et cocktail party (50%) pour différentes populations de sujets sourds, adapté de Turner et al., 2008.

Lorsqu'une courbe n'apparaît pas sur le graphe, c'est que la mesure n'a pas pu être réalisée chez le sujet.

1^{er} Patient

Les données relatives au patient 1 sont présentées dans la figure 16.

Ce patient avait 68 ans au moment de l'implantation, sa surdité est évolutive et elle est apparue vers l'âge de 50 ans soit une durée de surdité d'environ 18 ans.

Les audiogrammes pré- et post-implantation montrent que l'audition a été relativement bien conservée chez ce sujet. Nous constatons qu'il a également une bonne dynamique résiduelle.

La stimulation acoustique seule du côté implanté permet de stimuler les sons jusqu'à 1400 Hz. D'après la vocale filtrée nous voyons que, en post-implantation, ce patient utilise toujours les sons (acoustiques) jusqu'à 2 kHz pour comprendre. Ceci est sans doute lié à la dynamique relativement bonne jusqu'à 2 kHz.

Pour autant, son audiométrie vocale montre des scores plutôt médiocres qui n'évoluent pas avec le temps. Pour le test de compréhension dans le bruit, le seuil à 50% se situe vers un RSB = 3 ou 4 dB, ce qui est proche du résultat obtenu par l'implanté cochléaire moyen. Chez ce patient, il se produit probablement une redondance des informations acoustiques et électriques qui serait dû à un chevauchement fréquentiel au niveau des fréquences medium. Ceci, ainsi que la longue durée de surdité pourraient expliquer les mauvaises performances chez ce patient.

Avec le recul et d'après les résultats que nous avons trouvés dans nos différentes études, nous pensons que ce patient n'aurait pas du être implanté. Il aurait été peut-être plus judicieux de proposer à ce patient une aide auditive à compression fréquentielle, malheureusement cette technologie n'existait pas au moment où nous avons proposé l'implantation au patient.

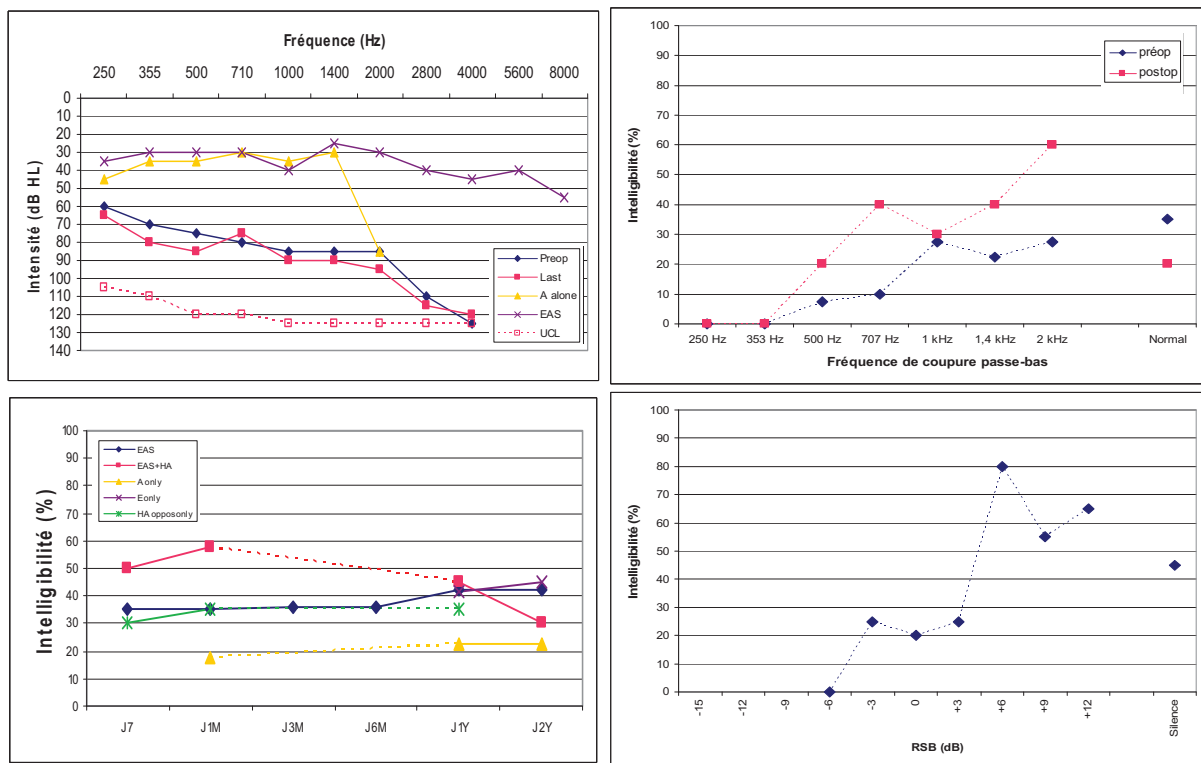


Figure 16. Données recueillies chez le 1^{er} sujet implanté EAS à Lyon. En haut à gauche : audiométries tonales ; en haut à droite : audiométrie vocale filtrée ; en bas à gauche : audiométrie vocale ; en bas à droite : audiométrie vocale dans le bruit.

2^{ème} Patient

Les données relatives au patient 2 sont présentées dans la figure 17.

Cette patiente avait 73 ans au moment de l'implantation, elle a eu une surdité brusque vers l'âge de 54 ans soit une durée de surdité d'environ 19 ans. La surdité est ensuite devenue évolutive. Cette patiente est la seule du groupe à avoir été implantée avec une électrode Flex Soft (longue).

Les audiogrammes pré- et post-implantation montrent que l'audition s'est dégradée d'environ 15-20 dB et la dynamique résiduelle est assez correcte. Cette patiente peut tout de même bénéficier de la stimulation acoustique sur les 700 premiers Hz. L'audiométrie vocale filtrée quant à elle n'a pas vraiment changé entre les mesures pré- et post-implantation.

Les résultats à l'audiométrie vocale montrent que la stimulation acoustique seule ne fournit aucune intelligibilité, mais la combinaison des 2 stimulations améliore les performances par rapport à l'implant seul. Chez cette patiente, le score d'intelligibilité ne dépasse jamais les 30% après 2 ans de rééducations. Ces performances sont sans doute liées à la longue durée de surdité chez cette patiente.

Nous pouvons nous demander si le choix d'implanter cette patiente était vraiment judicieux dans la mesure où elle présentait déjà un facteur limitant, à savoir une surdité ancienne. Mais ses aides auditives amplificatrices ne lui fournissaient aucun bénéfice (c'est d'ailleurs ce que nous retrouvons en post-implantation) et vu sa perte nous n'avions pas vraiment d'autre alternative que l'implantation. De plus l'audiométrie vocale filtrée pré-implantation nous orientait plutôt vers une implantation cochléaire standard, c'est pourquoi nous avons choisi d'utiliser une électrode longue chez cette patiente.

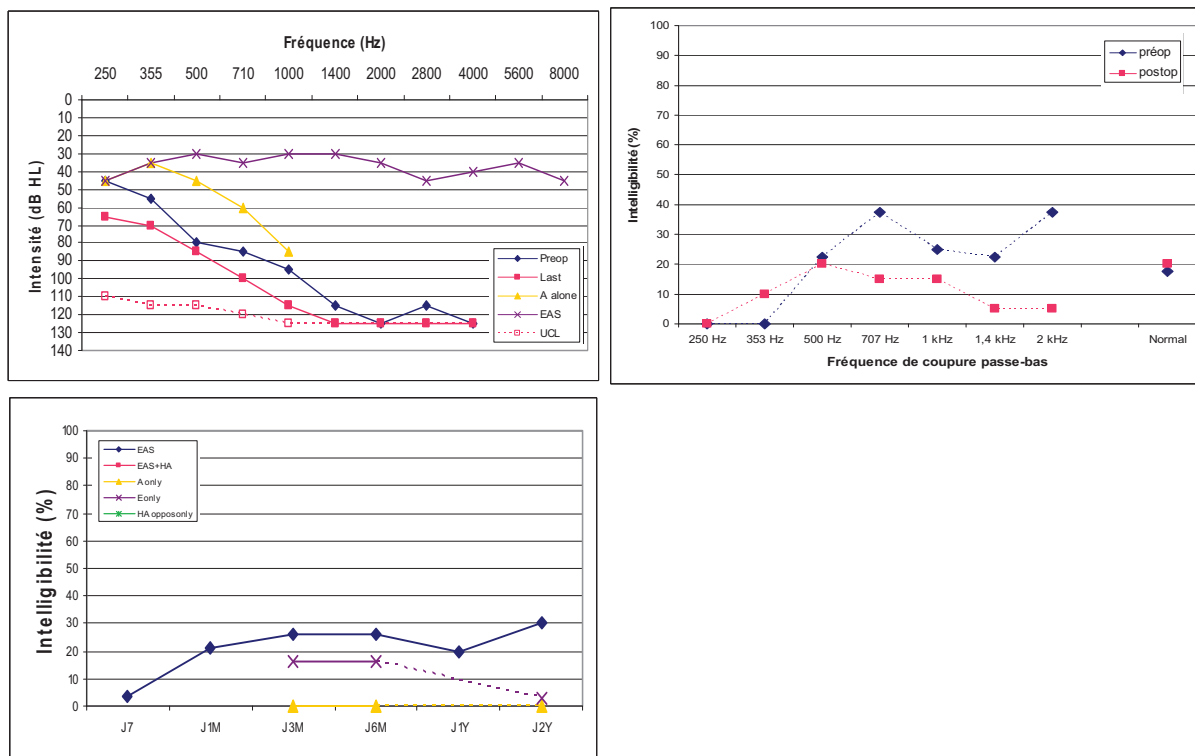


Figure 17. Données recueillies chez le 2^{ème} sujet implanté EAS à Lyon. En haut à gauche : audiométries tonales ; en haut à droite : audiométrie vocale filtrée ; en bas à gauche : audiométrie vocale.

3^{ème} Patient

Les données relatives au patient 3 sont présentées dans la figure 18.

Cette patiente avait 30 ans au moment de l'implantation, sa surdité est évolutive et elle est apparue vers l'âge de 15 ans soit une durée de surdité d'environ 15 ans.

Chez cette patiente l'audition a été relativement bien conservée après la chirurgie et la dynamique à J 2ans est très correcte. D'après l'audiométrie vocale filtrées pré-implantation, nous pensons que la patiente serait une bonne indication d'EAS, dans la mesure où la partie acoustique pourrait lui fournir 40-50% d'intelligibilité avec les 700-1000 premiers Hz et c'est ce que nous retrouvons effectivement à J 2ans.

Au niveau des résultats de compréhension, cette patiente a obtenu très vite de très bonnes performances. Le seuil de reconnaissance à 50% dans le bruit est obtenu à RSB = -6 dB environ, ce qui est bien meilleur que l'implanté cochléaire utilisateur d'une prothèse controlatérale.

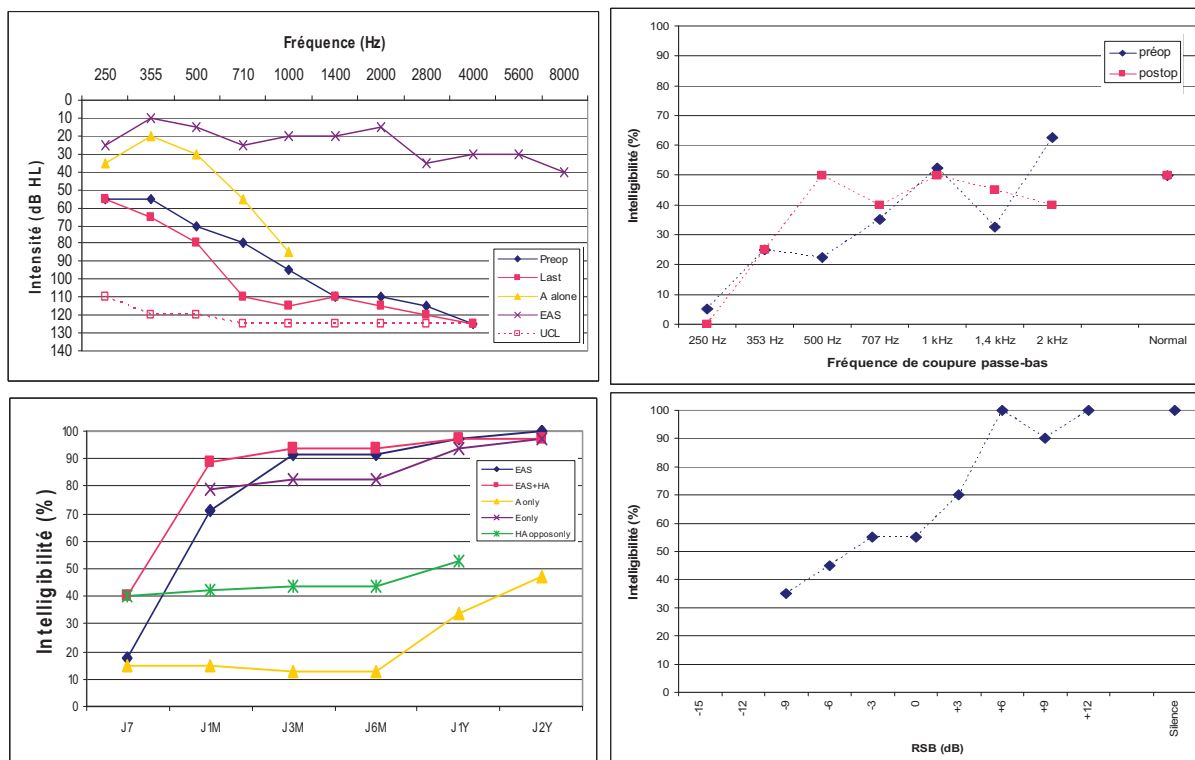


Figure 18. Données recueillies chez le 3^{ème} sujet implanté EAS à Lyon. En haut à gauche : audiométries tonales ; en haut à droite : audiométrie vocale filtrée ; en bas à gauche : audiométrie vocale ; en bas à droite : audiométrie vocale dans le bruit.

4^{ème} Patient

Les données relatives au patient 4 sont présentées dans la figure 19.

Cette patiente âgée de 37 ans au moment de l'implantation a eu une surdité brusque apparue à l'âge de 31 ans soit une durée de surdité d'environ 6 ans.

En terme de préservation de l'audition, cette patiente a vu ses seuils s'élever d'environ 20 dB après l'implantation avec une dynamique nulle au-delà de 500 Hz. La patiente a tout de même bénéficié de la stimulation acoustique pour les 700-1000 premiers Hz. D'après l'audiométrie vocale filtrée pré-implantation, nous pensions que la patiente était potentiellement une bonne indication d'EAS, dans la mesure où elle n'utilisait pas acoustiquement les sons au-delà de 500-700 Hz pour comprendre et dans la mesure où la partie acoustique pourrait lui fournir 30 à 40% d'intelligibilité.

En terme de performance, cette patiente a mis plusieurs mois avant de bien comprendre avec l'EAS et à J 2ans les performances sont très bonnes.

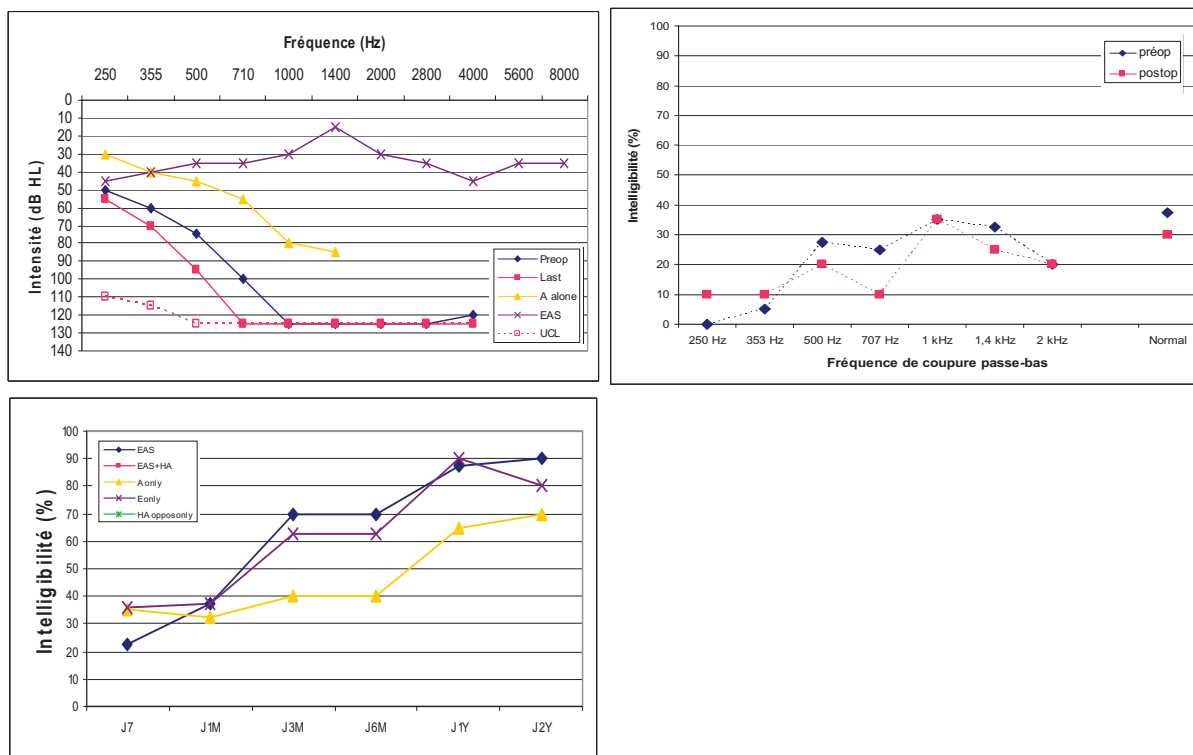


Figure 19. Données recueillies chez le 4^{ème} sujet implanté EAS à Lyon. En haut à gauche : audiométries tonales ; en haut à droite : audiométrie vocale filtrée ; en bas à gauche : audiométrie vocale.

5^{ème} Patient

Les données relatives au patient 5 sont présentées dans la figure 20.

Cette patiente âgée de 38 ans a une surdité qui est apparue dans la petite enfance mais qui évolue très lentement ; pour nos analyse nous avons considérée que sa durée de surdité dans les hautes fréquences était de 38 ans.

Cette patiente a vu son audition se dégrader de façon évolutive après l'implantation, et à J 2 ans la chute moyenne d'audition est d'environ 30-35 dB. Cette patiente était une bonne indication d'après notre test d'audiométrie vocale passe-bas. Malheureusement la patiente n'a pas pu bénéficier de la stimulation acoustique pour confirmer nos prédictions.

Cette patiente peut toutefois bénéficier de la stimulation bimodale grâce à son aide auditive controlatérale qui lui permet de très bien comprendre la parole dans le silence et dans le bruit. C'est d'ailleurs la patiente qui a obtenu la meilleure résistivité au bruit sur le groupe de 50 sujets implantés cochléaires que nous avons testés dans l'article 2. Pour cette patiente, le seuil de reconnaissance de la parole se situe entre environ -9 et -6 dB. (Le test dans le bruit a été réalisé lors d'une visite intermédiaire entre J 1 an et J 2 ans, c'est pourquoi le score de 100% obtenu dans le silence dans le cadre en bas à droite ne se retrouve pas sur le graphe en bas à gauche)

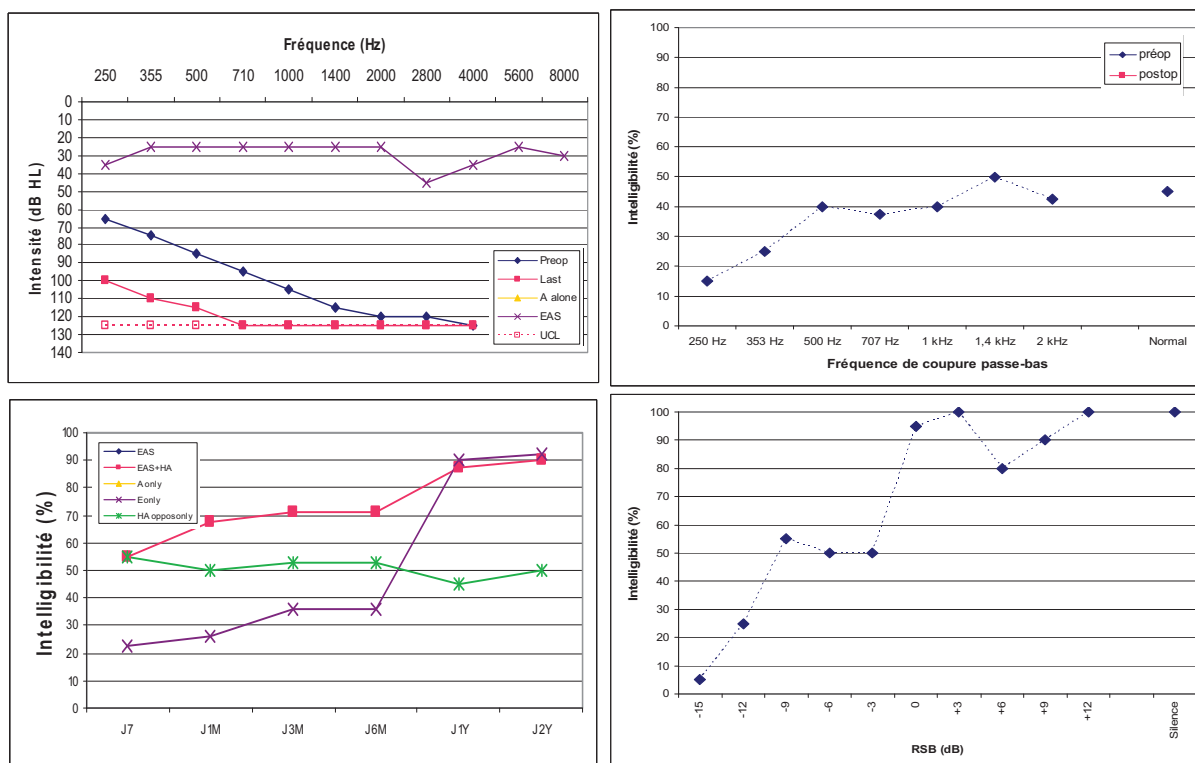


Figure 20. Données recueillies chez le 5^{ème} sujet implanté EAS à Lyon. En haut à gauche : audiométries tonales ; en haut à droite : audiométrie vocale filtrée ; en bas à gauche : audiométrie vocale ; en bas à droite : audiométrie vocale dans le bruit.

6^{ème} Patient

Les données relatives au patient 6 sont présentées dans la figure 21.

Ce patient avait 22 ans au moment de l'implantation, il présente une surdité lentement évolutive apparue dans la petite enfance soit une durée de privation dans les hautes fréquences d'environ 22 ans.

Ce patient a eu une très bonne conservation de son audition post-implantation, excepté pour le 250 Hz qui s'est dégradé d'environ 30 dB entre la mesure pré-implantation et J 2ans. Ce patient a gardé une dynamique auditive sur l'ensemble des fréquences testées. Il bénéficie de la stimulation acoustique, mais l'audiogramme avec la partie acoustique seule n'a pas pu être réalisé post-implantation.

D'après l'ensemble des tests audiométriques pré-implantatoires, nous avons prédit que ce patient pourrait être une bonne indication à l'EAS. Cependant le sujet ne porte pas son implant sur une base régulière, d'où une évolution très lente des performances et une compréhension autour de 10% avec l'EAS. Ceci est probablement dû à la surdité apparue dans l'enfance qui fait que le patient a du mal à « accepter » la stimulation de l'implant. En effet le sujet utilise surtout son aide auditive controlatérale. A J 2 ans, il parvient à obtenir une intelligibilité de l'ordre de 80% lorsqu'il utilise sa prothèse controlatérale et son implant.

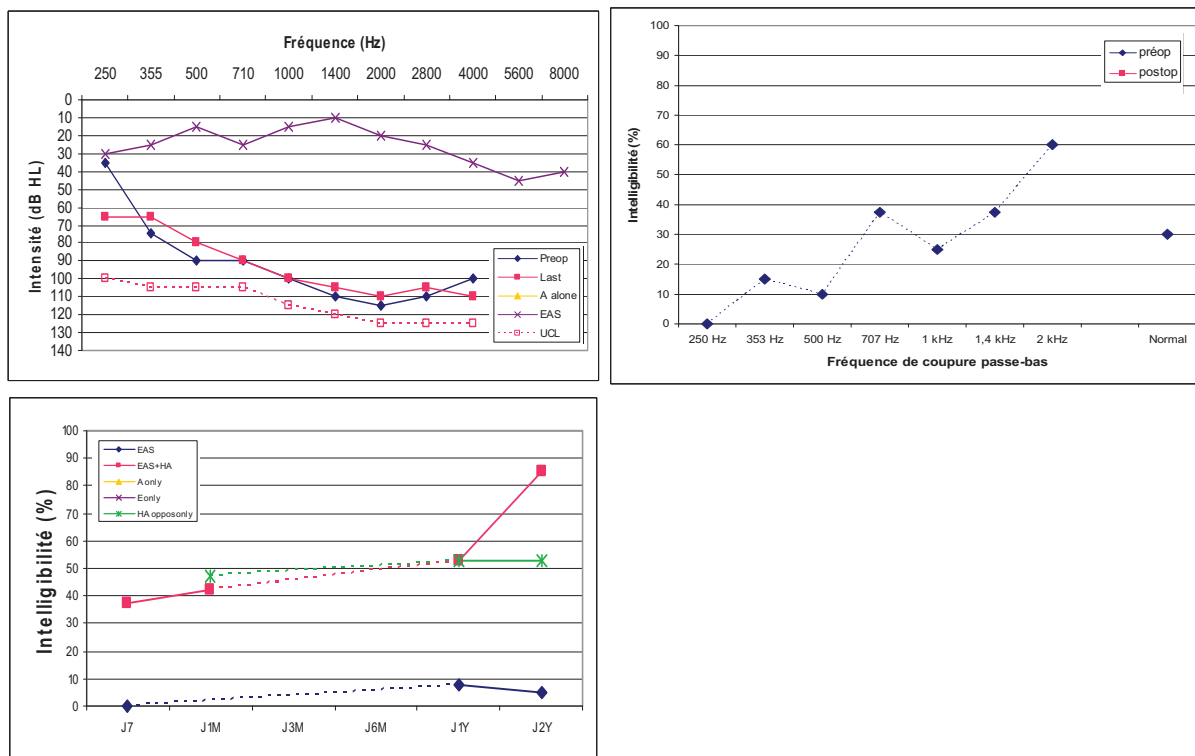


Figure 21. Données recueillies chez le 6^{ème} sujet implanté EAS à Lyon. En haut à gauche : audiométries tonales ; en haut à droite : audiométrie vocale filtrée ; en bas à gauche : audiométrie vocale.

7^{ème} Patient

Les données relatives au patient 7 sont présentées dans la figure 22.

Cette patiente est âgée de 36 ans, sa surdité est apparue de façon brusque pendant l'enfance. Nous avons considérée que la durée de surdité pré-implantation était 36 ans.

Cette patiente a eu une chute d'audition d'environ 40 dB entre les seuils obtenus pré-implantation et à J 2ans et sa dynamique est nulle au-delà de 350 Hz. Cependant la partie acoustique de l'implant lui permet de bénéficier des 700 premiers Hz à un niveau proche de son audition oreille nue pré-implantation. Du point de vue de l'audiométrie vocale filtrée, nous avons prédit que cette patiente pourrait être une bonne indication d'EAS puisqu'elle n'utilisait pas les sons au-delà de 1000 Hz pour comprendre et la stimulation acoustique des 700-1000 premiers Hz pouvait lui fournir 50% d'intelligibilité.

Tout comme la 4^{ème} implantée, cette patiente a mis plusieurs mois avant de bien comprendre avec l'implant EAS mais à J 1an, les performances de compréhension dans le silence et dans le bruit sont très bonnes, le seuil de compréhension dans le bruit se situe autour de 0 dB.

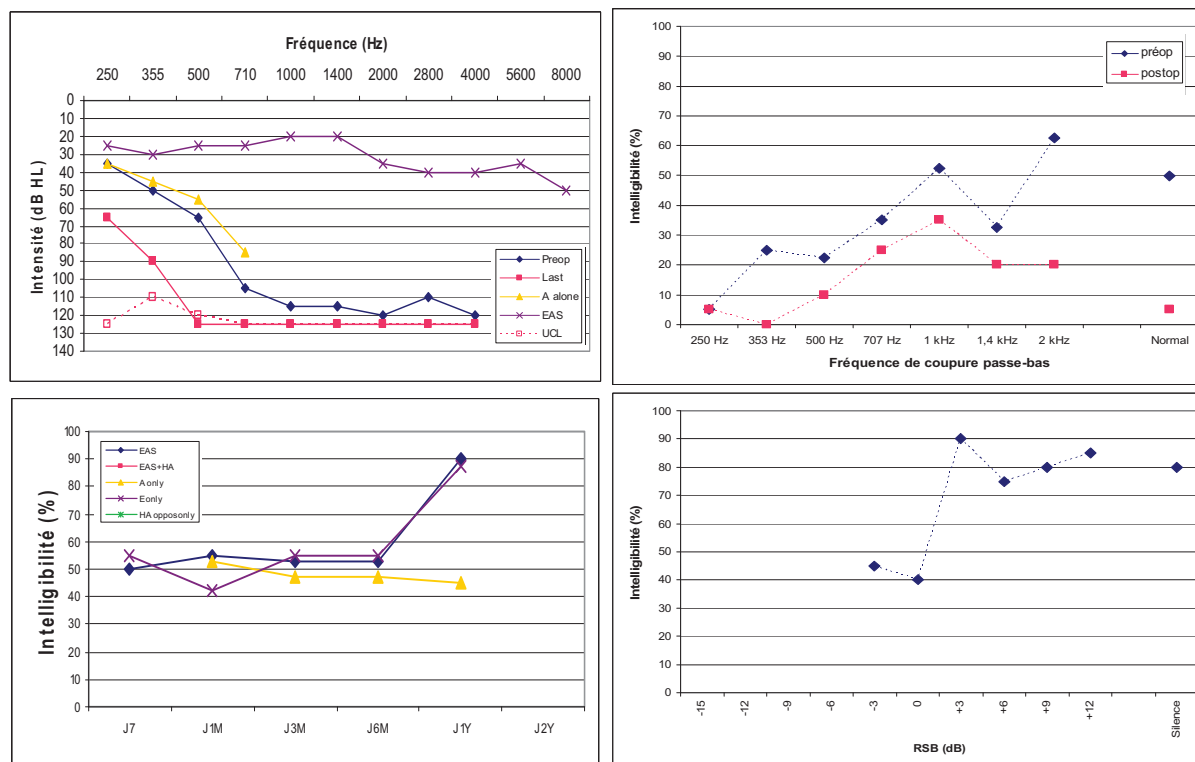


Figure 22. Données recueillies chez le 7^{ème} sujet implanté EAS à Lyon. En haut à gauche : audiométries tonales ; en haut à droite : audiométrie vocale filtrée ; en bas à gauche : audiométrie vocale ; en bas à droite : audiométrie vocale dans le bruit.

8^{ème} Patient

Les données relatives au patient 8 sont présentées dans la figure 23.

Ce patient âgé de 69 ans a une surdité évolutive qui est apparue vers l'âge de 60 ans, soit une durée de surdité de 9 ans.

Ce patient a vu son audition se dégrader progressivement après l'implantation, aussi bien en terme de seuil de perception qu'en terme de compréhension. Nous constatons qu'à J 2 ans sa dynamique est nulle au-delà de 350-500 Hz. D'après notre test d'audiométrie vocale filtrée, nous avons prédit que ce patient était une bonne indication d'EAS. Malheureusement le patient n'a pas pu bénéficier de la stimulation acoustique et nous n'avons pas pu vérifier nos prédictions.

En terme de compréhension, ce patient a mis plusieurs mois à « démarrer » mais les performances sont très fluctuantes d'une session de test à l'autre et elles ont même diminué entre J 1an et J 2ans. Pour la compréhension dans le bruit (implant + prothèse controlatérale), les performances sont très mauvaises aussi, le seuil de compréhension se situe probablement à un RSB entre +6 et +12 dB mais il n'a pas pu être évalué avec précision car le patient était vite fatigué par le test (le test dans le bruit a été réalisé lors d'une visite intermédiaire, c'est pourquoi le score de 100% obtenu dans le silence dans le cadre en bas à droite ne se retrouve pas sur le graphe en bas à gauche).

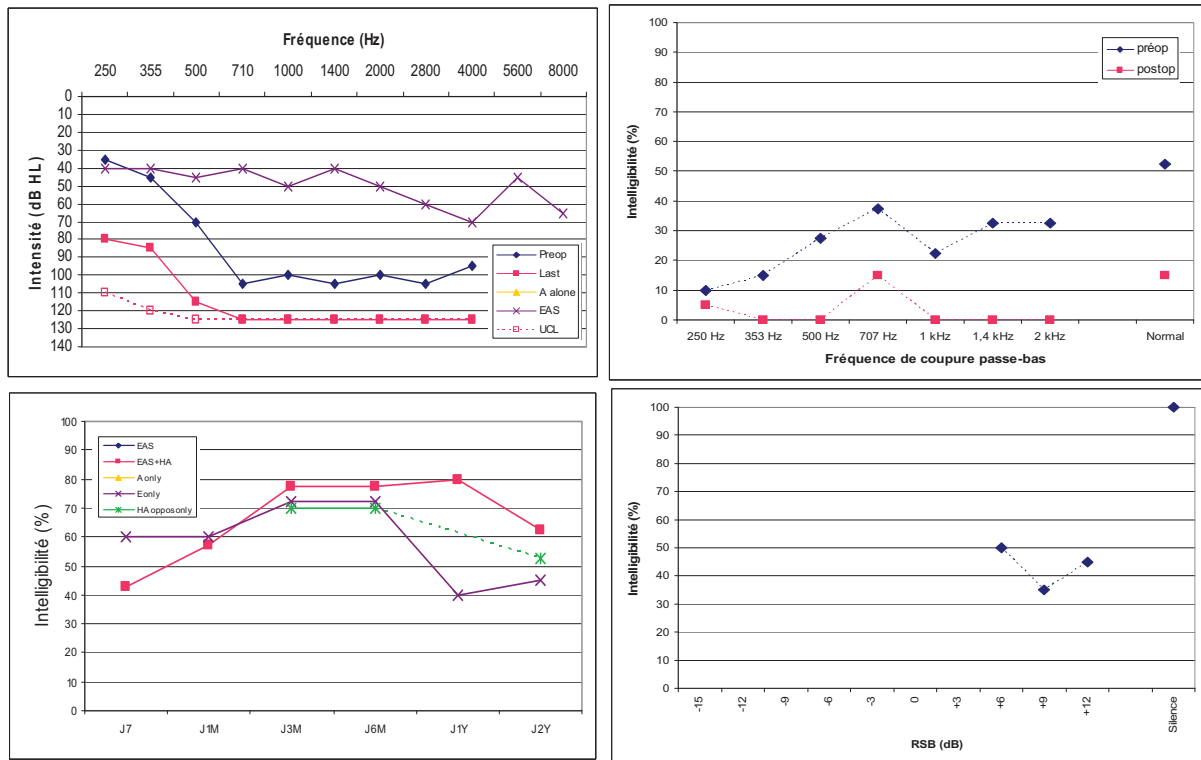


Figure 23. Données recueillies chez le 8^{ème} sujet implanté EAS à Lyon. En haut à gauche : audiométries tonales ; en haut à droite : audiométrie vocale filtrée ; en bas à gauche : audiométrie vocale ; en bas à droite : audiométrie vocale dans le bruit.

Bilan

Si nous faisons un bilan sur nos 8 patients implantés EAS, nous constatons que l'audition a été perdue chez 2/8 sujets. Quatre sujets ont de bonnes performances et 4 sujets ont des performances plutôt mauvaises. De manière globale pour la compréhension dans le bruit, sur 50 sujets implantés cochléaires testés, le groupe de sujets EAS est le groupe qui obtient les meilleures performances.

Le constat que la rééducation peut demander plus de temps chez certains patients EAS est en accord avec ce qui est décrit dans la littérature (Gantz et Turner, 2003).

Désormais il faut essayer de comprendre à quoi sont liés les mauvais résultats chez certains patients. La durée de privation joue sans doute un rôle majeur sur les résultats, d'ailleurs la majorité de nos patients ont une surdité relativement ancienne. Mais notre 5^{ème} patiente qui a une surdité ancienne de 38 ans est pourtant celle qui obtient les meilleures performances dans le bruit avec son implant et sa prothèse controlatérale. C'est donc qu'il y a d'autres facteurs qui interviennent. Il faudrait par exemple relier les performances à d'autres paramètres tels que la dynamique ou le résultat à l'audiométrie vocale filtrée, mais notre effectif de patients est actuellement trop faible pour pouvoir tirer des conclusions.

II.5. Simulation d'un implant EAS avec un vocoder hybride : Influence du nombre de canaux sur l'intelligibilité de la parole

(article publié : Seldran F, Thai-Van H, Truy E, Berger-Vachon C, Collet L, Gallego S, Beliaeff M. Simulation of an EAS implant with a hybrid vocoder. Cochlear Implants Int. 2010 Jun;11 Suppl 1:125-9)

Comme nous l'avons décrit dans l'introduction, les firmes d'implants cochléaires proposent pour l'EAS des porte-électrodes atraumatiques, souvent plus courts que les porte-électrodes standards, mais le nombre de canaux sur le faisceau implanté reste inchangé. Med-El propose par exemple le porte-électrode Flex EAS (24 mm ; longueur active de 20,9 mm) alors que le porte-électrode standard mesure 31,5 mm (longueur active de 26,4 mm) tous les deux possèdent 12 canaux de stimulation.

Par rapport aux faisceaux standards, les porte-électrodes conçus pour l'EAS ont donc des contacts (canaux) qui sont beaucoup plus proches les uns des autres. La réduction de cet espace inter-électrode de 2,4 à 1,9 mm représente une réduction de 20%. De plus, les surfaces des 5 contacts apicaux sont plus petites par rapport au porte-électrode standard. Ces deux changements cumulés (réduction de l'espace inter-électrode et réduction de la surface de stimulation) risquent de favoriser les interactions inter-canaux, ce qui peut dégrader la qualité du message électrique délivré au nerf auditif et conduire à une dégradation des performances des implantés.

Une solution envisageable pour réduire les interactions serait de réduire le nombre d'électrodes sur la partie implantée, ce qui permettrait d'espacer les contacts. Des travaux de la fin des années 90's ont mis en évidence qu'un nombre de canaux supérieur à 8 n'améliorait pas significativement les performances pour l'implant cochléaire. Par transposition à l'EAS, nous pensons que ce nombre peut être réduit si le patient présente des reliquats d'audition dans les basses fréquences.

L'objectif de cette étude est d'évaluer par un vocoder hybride, le nombre de canaux suffisants à la partie implantée pour fournir une bonne intelligibilité de la parole dans le calme. Si les résultats montrent qu'il suffit de peu de canaux pour fournir une bonne intelligibilité, il pourrait être intéressant pour le futur de l'EAS, de concevoir des implants mono- ou bi-canaux pour réhabiliter les hautes fréquences dans le cas de surdités partielles et ceci aurait l'avantage de minimiser grandement les interactions. L'autre avantage serait un nombre de fils réduits à l'intérieur du faisceau, ce qui pourrait probablement le rendre plus flexible et ainsi moins traumatique.

Notre simulation a été évaluée sur une population de 24 sujets normo-entendants. Nous avons reproduit la stimulation acoustique par un filtrage passe-bas à différentes fréquences de coupures pour reproduire différentes configurations de reliquats auditifs. L'implant quant à lui a été reproduit par une technique type vocoder avec 1, 2, 3, ou 4 canaux et une stratégie type CIS, à savoir que tous les canaux sont stimulés simultanément. Le codage de l'implant simulé se fait depuis la fréquence de coupure acoustique jusqu'à 4000 Hz. Les tests sont réalisés dans le calme.

Pour l'interprétation de nos résultats, nous avons considéré que les performances étaient satisfaisantes lorsqu'elles atteignent 90% ou plus. Avec ce critère, nos mesures ont montré que nous pouvions obtenir une bonne intelligibilité :

- avec 1 canal si l'audition résiduelle s'étend jusqu'à 1400 Hz ;
- avec 2 canaux si l'audition résiduelle s'étend jusqu'à 1000 Hz ;
- avec 3 canaux si l'audition résiduelle s'étend jusqu'à 700 Hz ;
- avec 4 canaux si l'audition résiduelle s'étend jusqu'à 500 Hz.

Ces résultats suggèrent que pour l'EAS, nous pourrions envisager d'utiliser des porte-électrodes avec un nombre de contacts plus espacés et plus réduits.

Simulation of an EAS Implant with a Hybrid Vocoder

FABIEN SELDRAN, HUNG THAI-VAN, ERIC TRUY,
CHRISTIAN BERGER-VACHON, LIONEL COLLET, and
STÉPHANE GALLEG0

*CNRS UMR 5020, Laboratoire Neurosciences Sensorielles, Comportement,
Cognition — Hôpital Edouard HERRIOT, Service d'audiologie et
d'explorations orofaciales, Lyon, France*

FABIEN SELDRAN AND MICHEL BELIAEFF

VIBRANT MED-EL Hearing Technology GmbH, Sophia-Antipolis, France

Introduction

Although EAS electrode arrays are shorter, they have the same number of electrodes as conventional implant arrays. The possibility of channel interaction already exists with a standard length electrode array. As electrodes on an EAS array are closer together, the likelihood of channel interaction is even greater. To compensate for this, we proposed to deactivate several channels on the electrode array, on the premise that the reduction in speech intelligibility due to fewer channels would be offset by the low frequency residual hearing. Recent studies have demonstrated that low frequency residual hearing provides significant benefit for speech understanding (Dorman et al., 2005; Turner et al., 2004). This study aims to determine the number of channels necessary to provide good speech understanding according to the degree of low frequency residual hearing, particularly the boundary or cutoff frequency (F_c) between acoustic and electrical stimulation. This study utilized an acoustic simulation of an EAS device in normal-hearing subjects.

Methods

Twenty-four young normal-hearing subjects participated in this study. They were tested through an earphone in the right ear without prior practice. The phonetic material used was the French Fournier word set (40 lists, each composed of 10 dissyllabic words) presented by a male speaker. Subjects were asked to repeat the word they had heard. The number of syllables repeated correctly was scored. As each list contains 10 dissyllabic words, each syllable repeated correctly was equal to 5 per cent recognition. Each list was played once and was presented in random order to prevent learning of the test material by the subjects. The tests were performed in quiet.

To simulate the auditory sensation perceived by an EAS recipient, a hybrid vocoder was designed using previously described methods (Shannon *et al.*, 1995). This computer program received acoustic input and produced an acoustic signal as presumably heard by an EAS recipient. This sound is then heard by normal hearing listeners for the experiment. Several vocoder parameters can be adjusted:

- 1) boundary between acoustic and electrical hearing
($F_c = 500$ Hz; 707 Hz; 1000 Hz; 1414 Hz)
- 2) type of stimulation (acoustic only; electric only; EAS = both acoustic+electric)
- 3) number of channels on the simulated implant (1 to 4 channels).

The 'Cool Edit Pro' software was used for signal processing. This software allows the user to design steeply sloping filters. The 'FFT Filter' function was used, with a coefficient of 100 per cent for the bandwidth and 0 per cent for the rejected band which reduces all spectral coefficients of the attenuated band by 70 dB.

The first simulation condition evaluated was 'acoustic only', i.e., the patient's residual low-frequency hearing amplified with a hearing aid. Low-pass filtering of the signal was performed up to the cut-off frequency ($F_c = 500$, 707, 1000, or 1414 Hz). Low-pass filtering was sufficient to essentially simulate the 'frequency limitation' of high frequency hearing impaired patients.

The second simulation condition assessed was 'electric only', i.e., only the EAS electrode array. Most of the spectral components of speech are contained in frequencies below 4 kHz. Therefore, simulation of the 'electric only' condition was limited to frequencies in the F_c to 4 kHz range. As the electrode array encodes high-frequency signals, channels were implemented based on a logarithmic scale in order to more closely reflect the tonotopic organization of the cochlea.

The F_c to 4 kHz range was then divided into N bands; N corresponding to the number of channels on the electrode array and equal to 1, 2, 3, or 4. For each channel, the bandwidth was limited by a band-pass filter (between F_{min} and F_{max}), and its envelope was extracted by full-wave rectification and low-pass filtering until 50 Hz. Every envelope was then filled with white noise, filtered with the same frequency range, but with a narrower bandwidth. Each envelope with a bandwidth between F_{min} and F_{max} was filled with white noise filtered with a band-pass between $(F_{min} + 75\text{Hz})$ and $(F_{max} - 75\text{Hz})$. Each channel was then amplified to compensate for the lack of energy induced by the white noise filling the envelope. It was ensured that the energy contained in the simulated electrical channel was equal to the energy contained in the initial acoustic signal, filtered by the band-pass corresponding to that channel.

The third simulation condition corresponded to EAS hearing. To simulate the electrical plus acoustical stimulation as experienced by EAS patients, the signals obtained in the two previous conditions ('acoustic only' and 'electric only') were combined. Figure 1 illustrates the spectral resolution of the acoustic channel of the signal ($F_c = 500$ Hz) compared to the four electrical channels.

Results

Speech understanding was deemed to be satisfactory when syllable intelligibility scores were over 90 per cent. In order to restore speech understanding in quiet (Figure 2):

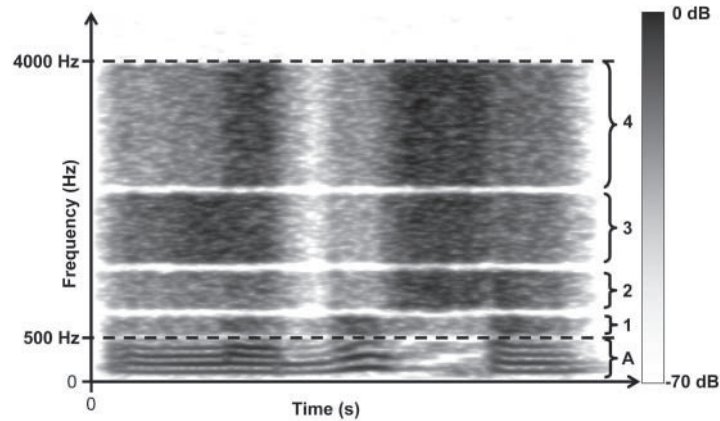


FIGURE 1 Spectrogram of the French word «le bouchon» encoded for the condition $F_c=500$ Hz with 4 electrical channels. The brackets show the different bands:
A=ACOUSTIC stimulation; 1, 2, 3, and 4=N, number of simulated electrical channels

1. electrical channel (■) is sufficient for a patient with residual hearing up to 1400 Hz
2. channels (▲) are sufficient for a patient with residual hearing up to 1000 Hz
3. channels (●) are necessary for a patient with residual hearing up to 700 Hz
4. channels (×) are necessary for a patient with residual hearing up to 500 Hz.

Current audiometric EAS candidacy criteria are sufficient residual hearing to 700–1000 Hz and a speech recognition score of approximately 60 per cent. With acoustic stimulation only (◆) until 700 or 1000 Hz, syllable intelligibility scores of 55–65 per cent may be attained (Figure 2). This syllable intelligibility result corresponds to the EAS candidacy criteria recommended by the manufacturer. With electrical stimulation only, it is possible to understand no more than approximately 45 per cent of syllables even with 4 channels (Figure 3).

Discussion

Previous work has documented speech recognition scores of 0–13 per cent for monosyllabic words with acoustic stimulation only, 45–61 per cent with electrical stimulation only, and 50–75 per cent for the EAS (acoustic+electric) condition (von Ilberg et al., 1999; Kiefer et al., 2005; Gstöettner et al., 2006). These results are consistent with our findings for both the EAS and ‘electrical only’ conditions. It is difficult to compare results for the ‘acoustic only’ condition as the degree of residual hearing was not described in these studies. The performance in the EAS condition was superior to the sum of the performance measured in both the ‘acoustic only’ and the ‘electric only’ conditions. The combination of both electric and acoustic stimulation produced a synergistic effect.

We plan to reproduce this study in noisy surroundings to simulate everyday life. Further work will also include the evaluation of speech perception in EAS patients, in the same three conditions (acoustic only, electric only and EAS).

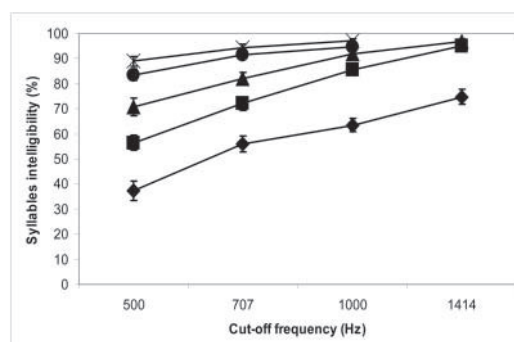


FIGURE 2 Syllable intelligibility (%) in quiet versus cutoff frequency between acoustic and electric stimulation (mean results for 24 normal-hearing people).

Curves represent (from bottom to top):

- acoustic stimulation only (◆)
- acoustic stimulation+1 electric channel (■)
- acoustic stimulation+2 electric channels (▲)
- acoustic stimulation+3 electric channels (●)
- acoustic stimulation+4 electric channels (×)

Satisfactory speech understanding (syllable intelligibility scores > 90%) with:

- 1 electric channel if residual hearing is up to 1400 Hz
- 2 electric channels if residual hearing is up to 1000 Hz
- 3 electric channels if residual hearing is up to 700 Hz
- 4 electric channels if residual hearing is up to 500 Hz.

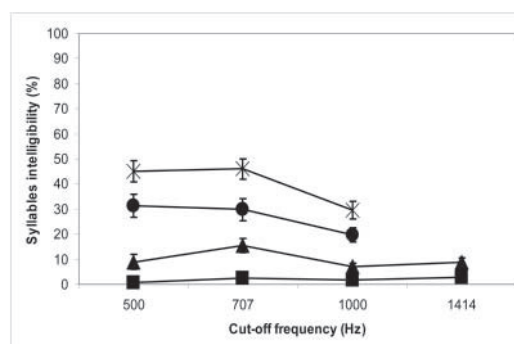


FIGURE 3 Syllable intelligibility (%) in quiet versus cutoff frequency between acoustic and electric stimulation for 'electric only' condition (mean results for 24 normal-hearing people).

Curves represent (from bottom to top):

- 1 electric channel (■)
- 2 electric channels (▲)
- 3 electric channels (●)
- 4 electric channels (×)

Conclusion

This simulation study demonstrated that 4 channels on the EAS electrode array are sufficient to provide good speech perception. Reducing the number of electrode channels does not impair speech understanding of patients with residual hearing. It is hypothesized that a reduction in the number of channels will result in less electrical interaction between electrodes leading to better performance. To assess factors important to better speech perception, additional simulation studies are planned implementing different coding strategies. This research may facilitate the development of new EAS electrode array designs with fewer channels.

Bibliography

- Dorman, M.F., Spahr, A.J., Loizou, P.C., Dana, C.J. & Schmidt, J.S. 2005. Acoustic Simulations of Combined Electric and Acoustic Hearing (EAS). *Ear Hear*, 26(4): 371–80.
- Gstoettner, W.K., Helbig, S., Maier, N., Kiefer, J., Radeloff, A. & Adunka, O.F. 2006. Ipsilateral Electric Acoustic Stimulation of the Auditory System: Results of Long-Term Hearing Preservation. *Audiol Neurotol*, 11(Suppl 1): 49–56. Epub 6 October 2006.
- Kiefer, J., Pok, M., Adunka, O., Stürzebecher, E., Baumgartner, W., Schmidt, M., Tillein, J., Ye, Q. & Gstoettner, W. 2005 May-Jun. Combined Electric and Acoustic Stimulation of the Auditory System: Results of a Clinical Study. *Audiol Neurotol*, 10(3): 134–44. Epub 17 February 2005.
- Shannon, R.V., Zeng, F.G., Kamath, V., Wygonski, J. & Ekelid, M. 1995. Speech Recognition with Primarily Temporal Cues. *Science*, 270(5234): 303–04.
- Turner, C.W., Gantz, B.J., Vidal, C., Behrens, A. & Henry, B.A. 2004. Speech Recognition in Noise for Cochlear Implant Listeners: Benefits of Residual Acoustic Hearing. *J Acoust Soc Am*, 115(4): 1729–35.
- von Ilberg, C., Kiefer, J., Tillein, J., Pfenningdorff, T., Hartmann, R., Stürzebecher, E. & Klinke, R. 1999. Electric-Acoustic Stimulation of the Auditory System. New Technology for Severe Hearing Loss. *ORL J Otorhinolaryngol Relat Spec*, 61(6): 334–40.

Correspondence to: fseldran@yahoo.fr

II.6. Analyse de « l'avantage de la stimulation combinée » par des modèles de combinaison de l'information

(article publié : Seldran F, Micheyl C, Truy E, Berger-Vachon C, Thai-Van H, Gallego S. A model-based analysis of the "combined-stimulation advantage" Hear Res. 2011 Dec;282(1-2):252-64)

Ces dernières années, les firmes d'implants cochléaires et les équipes de recherche portent un intérêt tout particulier à la stimulation bimodale et à l'amélioration des performances qu'elle fournit par rapport à la stimulation de l'implant cochléaire seule. La question qui se pose est « comment expliquer cette amélioration des performances ? ».

Kong et Carlyon (2007) ont montré qu'il se produisait un mécanisme de « super-additivité » qui résulterait d'interactions constructives lorsque les stimulations de l'implant et de la prothèse acoustique sont combinées. Nous pensons que l'explication de Kong et Carlyon est en réalité biaisée par l'utilisation d'un modèle qui n'est pas optimal (le modèle de sommation de probabilité) et que modèles alternatifs permettraient de mieux expliquer les performances observées en condition bimodale.

Dans cette étude, nous avons soumis 24 sujets normo-entendant à un test d'écoute de signaux de parole filtrés passe-bas, vocodés, et combinés (filtrés passe-bas + vocodés), dans le silence et à différents rapports signal/bruit, pour différentes fréquences de coupures acoustique/électrique, et pour différents nombres de canaux sur l'implant simulé. Différents modèles de combinaison des informations ont été testés pour expliquer nos données observées.

Les modèles utilisés sont tout à fait capables de prédire le score en condition bimodale, sur la base de ce qui est perçu en filtré passe-bas seul et en vocodé seul. Les modèles que nous avons utilisés n'impliquant pas de mécanismes d'interactions constructives, c'est donc que les bénéfices en condition bimodale simulée peuvent être expliqués sans invoquer des mécanismes d'interactions constructives dans la perception simultanée des stimulations vocodées et filtrées passe-bas.



Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research paper

A model-based analysis of the “combined-stimulation advantage”

Fabien Seldran^{a,b,c,d,*}, Christophe Micheyl^e, Eric Truy^{a,b,c,f}, Christian Berger-Vachon^{a,b,c}, Hung Thai-Van^{a,b,c,f}, Stéphane Gallego^{a,b,c}^a INSERM U1028, Lyon Neuroscience Research Center, PACS Team (Speech, Audiology, Communication Health), Lyon F-69000, France^b CNRS UMR5292, Lyon Neuroscience Research Center, PACS Team (Speech, Audiology, Communication Health), Lyon F-69000, France^c University Lyon 1, Lyon F-69000, France^d Vibrant Med-El Hearing Technology GmbH, 400 Ave. Roumanille, BP 309, Sophia-Antipolis 06906, France^e Department of Psychology, University of Minnesota, Minneapolis, MN 55455, USA^f Audiology and ENT Department, Edouard Herriot Hospital, Lyon F-69437, France

ARTICLE INFO

Article history:

Received 20 May 2011

Received in revised form

19 June 2011

Accepted 20 June 2011

Available online 22 July 2011

ABSTRACT

Improvements in speech-recognition performance resulting from the addition of low-frequency information to electric (or vocoded) signals have attracted considerable interest in recent years. An important question is whether these improvements reflect a form of constructive perceptual interaction—whereby acoustic cues enhance the perception of electric or vocoded signals—or whether they can be explained without assuming any interaction. To address this question, speech-recognition performance was measured in 24 normal-hearing listeners using lowpass-filtered, vocoded, and “combined” (lowpass + vocoded) words presented either in quiet or in a realistic background (cafeteria noise), for different signal-to-noise ratios, different lowpass-filter cutoff frequencies, and different numbers of vocoder bands. The results of these measures were then compared to the predictions of three models of cue combination, including a “probability-summation” model and two Gaussian signal detection theory (SDT) models—one (the “independent-noises” model) involving pre-combination noises, and the other (the “late-noise” model) involving post-combination noise. Consistent with previous findings, speech-recognition performance with combined stimulation was significantly higher than performance with vocoded or lowpass stimuli alone, and it was also higher than predicted by the probability-summation model. The two Gaussian-SDT models could account quantitatively for the data. Moreover, a Bayesian model-comparison procedure demonstrated that, given the data, these two models were far more likely than the probability-summation model. Since these models do not involve any constructive-interaction mechanism, this demonstrates that constructive interactions are not needed to explain the combined-stimulation benefits measured in this study. It will be important for future studies to investigate whether this conclusion generalizes to other test conditions, including real EAS, and to further test the assumptions of these different models of the combined-stimulation advantage.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

A recent development in the field of cochlear implants (CIs) relates to “electro-acoustic stimulation” (EAS). EAS involves the

combination of electric stimulation (via a CI) and acoustic stimulation (usually, via a hearing aid), either in the same ear (e.g., Gantz and Turner, 2003; Turner et al., 2004) or in the opposite ear (e.g., Ching et al., 2004; Kong et al., 2005; Mok et al., 2006). Several studies have demonstrated improved speech-recognition performance with EAS, compared to electric or acoustic stimulation alone (e.g., Büchner et al., 2009; Cullington and Zeng, 2011; Gantz and Turner, 2003, 2004; Gantz et al., 2004, 2006; Gfeller et al., 2006; Turner et al., 2004; von Ilberg et al., 1999). Benefits of “combined-stimulation” have also been observed in “simulated-EAS” studies in normal-hearing listeners, in which CI processing was simulated using a vocoder, and residual low-frequency hearing was simulated by lowpass-filtering speech signals (e.g., Başkent and Chatterjee, 2010; Brown and Bacon, 2009a; Chang et al., 2006; Chen and Loizou, 2010; Dorman et al., 2005; Kong

Abbreviations: CF, cutoff frequency; CI, cochlear implant; EAS, electro-acoustic stimulation; F0, fundamental-frequency; PC, percentage of correct responses; RMS, root-mean-square; SDT, signal detection theory; SNR, signal-to-noise ratio.

* Corresponding author. INSERM U1028, CNRS UMR5292, Centre de Recherche en Neurosciences de Lyon, Equipe Audition, Hôpital Edouard Herriot, Pavillon U, Place d'Arsonval, F-69437 Lyon Cedex 03, France. Tel.: +33 4 72 11 05 03; fax: +33 4 72 11 05 04.

E-mail addresses: fseldran@yahoo.fr (F. Seldran), cmicheyl@umn.edu (C. Micheyl), eric.truy@chu-lyon.fr (E. Truy), christian.berger-vachon@univ-lyon1.fr (C. Berger-Vachon), hthaiivan@gmail.com (H. Thai-Van), sgallego@hotmail.fr (S. Gallego).

and Carlyon, 2007; Li and Loizou, 2008; Qin and Oxenham, 2006; Turner et al., 2004).

Although the potential and actual benefits of combining acoustic stimulation with electric stimulation are well demonstrated, it is still not entirely clear what explains these benefits. It has been suggested that the provision of fundamental-frequency (F0) information at low frequencies plays a key role (e.g., Başkent and Chatterjee, 2010; Brown and Bacon, 2009a,b, 2010; Kong et al., 2005; Qin and Oxenham, 2006; Turner et al., 2004). F0 differences between talkers (e.g., female and male) provide a cue for the perceptual separation of concurrent voices (Brokx and Nooteboom, 1982), and it has been suggested that this cue is more salient at low frequencies (Carlyon, 1996; Culling and Darwin, 1993; Qin and Oxenham, 2006). Low-frequency F0 cues could help listeners track and extract a voice among other voices in a vocoded or electric mixture. According to another explanation, low-frequency cues could facilitate “glimpsing,” i.e., selective listening to the target voice through “dips” in the temporal envelope or spectrum of the masker (Kong and Carlyon, 2007; Li and Loizou, 2008).

While these various explanations continue to be actively investigated, a more basic question, which remains without a clear answer, is whether the perceptual benefits of combined stimulation can be explained simply in terms of non-interactive cue combination, or whether they necessarily require to assume the existence of synergetic interactions in the perceptual processing of low-frequency and electric (or vocoded) speech cues—so that access to the former somehow facilitates or enhances the use of the latter. To gain clarity on this important issue, Kong and Carlyon (2007) compared listeners' speech-recognition performance in combined-stimulation conditions to predictions derived (using a model described in Section 2.4.1) under the hypothesis that the low-frequency and vocoded signals were identified independently, and that the identification decisions were then combined in an “additive” fashion. The results showed that listeners' performance was higher than predicted by the model. Kong and Carlyon (2007) interpreted this result as evidence for “super-additive” effects (i.e., constructive interactions) in the combination of information across lowpass- and vocoded stimuli. Chang et al. (2006) also noted that the performance of their listeners in combined-stimulation conditions was higher than predicted by adding the proportions of correct responses for the lowpass-alone and vocoded-alone stimuli. Earlier, Kong et al. (2005) pointed out that low-frequency acoustic signals that yielded essentially zero percent correct when presented alone, nonetheless enhanced speech-recognition performance when they were added to electric stimulation in the opposite ear of CI listeners. These results have been interpreted as evidence for the existence of significant synergetic interactions in the processing of lowpass- and vocoded stimuli, and have inspired the search for specific mechanisms (such as F0-guided segregation or glimpsing), which might explain these effects. However, two recent studies, one by Kong and Braidă (2011), the other by Michey and Oxenham (submitted for publication), indicate that it may not be necessary to posit interactions to explain the benefits of combined stimulation. Kong and Braidă (2011) found that listeners' speech-recognition performance in real and simulated-EAS conditions could be accounted for, to a large extent, using a non-interactive cue-combination model based on Gaussian signal detection theory (SDT). Consistent with this conclusion, Michey and Oxenham (submitted for publication) reanalyzed the data of previous simulated-EAS studies using Gaussian-SDT models, and they found that such models could explain the performance of the listeners in those studies without the need to assume the existence of interactions between low-frequency and vocoded cues.

The aim of this study was to investigate further the hypothesis that listeners' ability to combine acoustic speech cues at low frequencies with temporal envelope and degraded spectral cues contained in vocoded signals at higher frequencies, can be predicted using relatively simple psychophysical models of cue combination, which do not posit any interaction in the processing of the two types of cues. To this aim, speech-recognition performance was measured using lowpass-filtered, vocoded, and combined (lowpass-filtered + vocoded) signals under a variety of conditions — including different signal-to-noise ratios (SNRs), different lowpass-filter cutoff frequencies, and different numbers of vocoder bands — in 24 normal-hearing listeners. The results of these measures were then compared to predictions obtained using three models of cue combination, including the model used by Kong and Carlyon (2007), and two Gaussian-SDT models.

2. Methods

2.1. Subjects

Twenty-four subjects (11 females, 13 males, aged 19–35 years, mean = 23.6 years) took part in the study. All had normal hearing, defined as hearing thresholds of 20 dB HL or better at octave frequencies between 250 and 8000 Hz. In accord with the Declaration of Helsinki, written informed consent was obtained from all subjects prior to their inclusion in the study.

2.2. Stimuli and procedure

The speech stimuli consisted of 40 lists of 10 disyllabic words (Fournier, 1951) spoken by a single male talker and recorded on a compact disc (CD, 44.1-kHz sampling rate, with 16-bit quantization range). The “lowpass” (L) stimuli were produced by filtering these signals digitally (using the FFT-filter function of the Adobe Audition software) below a cutoff frequency (CF) of 500, 707, 1000, or 1414 Hz (stopband attenuation > 70 dB). The vocoded (V) stimuli were produced by, firstly, bandpass-filtering the original speech signals into $N = 1, 2, 3$, or 4 frequency bands, which are hereafter referred to as “analysis bands.” The cutoff frequencies of the analysis bands are listed in Table 1. For the highest CF (1414 Hz), only two N conditions were tested: $N = 1$ and $N = 2$. The lower frequency limit of the lowest analysis band was equal to CF. The upper frequency limit of the highest analysis band was fixed at 4000 Hz. The center frequencies of the analysis bands were equidistant on a logarithmic scale between CF and 4000 Hz. The temporal envelopes of the signals at the output of these analysis bands were then extracted using full-wave rectification, followed by lowpass-filtering at 50 Hz to eliminate pitch-related envelope fluctuations (Carroll and Zeng, 2007). Thirdly, the resulting temporal envelopes were used to modulate the amplitude of noise bands (“synthesis bands”). To avoid interactions between modulation sidebands, each synthesis band was 150-Hz (75-Hz on each side) narrower than the corresponding analysis band (Carroll and Zeng, 2007). Finally, the amplitude-modulated synthesis bands were scaled to have the same RMS amplitude as the corresponding analysis bands in the original signal, then they were summed to produce the final V signal. The “combined” (C) stimuli were produced by adding together L and V signals corresponding to the same CF.

Stimuli were either presented in quiet, or added to recordings of cafeteria noise. The latter were chosen to provide a realistic, everyday-life noise background. They contained unintelligible chatter from a large number of talkers, and occasional noises of clanging dishes or chairs being moved, typical of a cafeteria. The noise was processed in the same way as the speech. The processed speech and noise signals were recorded separately on the left and

Table 1

Low- and high-frequency limits of the vocoder synthesis bands for the different CF and *N* conditions.

Condition	Band number	Synthesis band limits (Hz)
CF = 500 Hz		
<i>N</i> = 1	1	575–3925
<i>N</i> = 2	1	575–1339
	2	1489–3925
<i>N</i> = 3	1	575–925
	2	1075–1925
	3	2075–3925
<i>N</i> = 4	1	575–766
	2	916–1339
	3	1489–2303
	4	2453–3925
CF = 707 Hz		
<i>N</i> = 1	1	782–3925
<i>N</i> = 2	1	782–1607
	2	1757–3925
<i>N</i> = 3	1	782–1185
	2	1335–2170
	3	2320–3925
<i>N</i> = 4	1	782–1015
	2	1165–1606
	3	1756–2517
	4	2667–3925
CF = 1000 Hz		
<i>N</i> = 1	1	1075–3925
<i>N</i> = 2	1	1075–1925
	2	2075–3925
<i>N</i> = 3	1	1075–1512
	2	1662–2444
	3	2594–3925
<i>N</i> = 4	1	1075–1339
	2	1489–1925
	3	2075–2753
	4	2903–3925
CF = 1414 Hz		
<i>N</i> = 1	1	1489–3925
<i>N</i> = 2	1	1489–2303
	2	2453–3925

right tracks of a CD. During the tests, the signals from the left and right tracks were mixed at different signal-to-noise ratios (SNRs): –6 dB, 0 dB, and +6 dB. For the Quiet condition (which may be thought of as corresponding to an infinitely large physical SNR, and is hereafter included as a fourth SNR condition), the cafeteria noise was turned off.

Together, three presentation modes (L, V, and C), four CFs, and four SNRs would have resulted in a total of 48 stimulus conditions. However, because it was anticipated that performance in the quiet C conditions with *N* = 3 or 4 and a CF of 1000 Hz would be near ceiling (100% correct), the 1414-Hz CF was tested only when *N* was less than 3. As a result, the number of conditions tested was equal to 32. For each condition, and each listener, one list of 10 dissyllabic words was drawn at random (without replacement) from the 40 lists of dissyllabic words. The 10 words were presented to the listener's right ear. Listeners were instructed to repeat each word after its presentation. For scoring, the total number of correctly repeated syllables in each test condition was divided by the number of syllables (20) and multiplied by 100 to yield a percent-correct (PC) score. The tests took approximately 3 h per listener. Testing was divided into two sessions of approximately 90 min each, on different days. The listeners did not receive training prior to the tests.

2.3. Apparatus

The stimuli were played on a CD player (PHILIPS – CD723) connected to an audiometer (MADSEN – Orbiter 922), and

delivered through TDH-39 earphones. Tests took place in a sound-proof booth at Edouard Herriot Hospital in Lyon. The study was approved by the local Ethics Committee (CPP Sud-Est IV, Centre Leon Berard de Lyon, France, N° ID RCB: 2008-A01479-46).

2.4. Models

2.4.1. The probability-summation model

The first model that was considered in this study has been used in various other contexts in the psychophysical literature, where it is sometimes referred to as the “probability-summation,” or “independent decisions,” model (for introductions, see: Green and Swets, 1966; Macmillan and Creelman, 2005; Treisman, 1998). The model has been used to predict the detection of simple events (e.g., signal versus no-signal) at the outputs of two or more sensory channels (e.g., Pelli, 1985; Pirenne, 1943). Fletcher (1953) used this model to predict the proportion of correct responses for the recognition of simultaneously presented bands of speech based on the intelligibility of each band in isolation (see also: Braida, 1991; Ronan et al., 2004). Boothroyd and Nitttrouer also used, and expanded, the model to analyze context effects in speech recognition (Boothroyd and Nitttrouer, 1988; Nitttrouer and Boothroyd, 1990). The latter work prompted the use of the model in a simulated-EAS study by Kong and Carlyon (2007). In this context, the probability of a correct response in combined-stimulation conditions is predicted as the complement of the product of the error probabilities in the corresponding non-combined-stimulation conditions, as follows.

$$P_{C,PS} = 1 - (1 - P_L)(1 - P_V). \quad (1)$$

In this equation, and the ones that follow, P_L denotes the probability of a correct response in the L condition, P_V denotes the probability of a correct response in the V condition, and P_C denotes the predicted probability of a correct response in the C condition; the subscript “_PS” was added to distinguish the predictions of the probability-summation model from the predictions of other models.

2.4.2. Gaussian-SDT models

The model described in the previous section falls in the category of “post-labeling” models (Braida, 1991; Ronan et al., 2004). These models assume that listeners, first, identify speech items (phonemes, syllables, or words) within each channel, then, combine the resulting identification decisions in some way, e.g., by selecting one of the two answers with a certain probability. If the observations on which the decisions are based are not dichotomous, this decision strategy is sub-optimal, meaning that it is possible to achieve higher performance than predicted by these models, simply, by using a different decision strategy. In general, optimal (maximum-likelihood) decision strategies involve a combination of continuous decision variables, or observations, across channels *before* a decision is made as to which item was presented (Green and Swets, 1966; Wickens, 2001). Models employing this type of decision strategy are known as “pre-labeling” models (Braida, 1991; Ronan et al., 2004). Usually, the pieces of information that are (assumed to be) combined in these models are continuous quantities, which are related to likelihoods or likelihood ratios—for instance, the likelihood that speech item *x* was presented, given that input y_i was received in channel *i*. Moreover, these models usually assume that listeners' ability to correctly identify speech signals is limited by variability, or “noise.” The “noise” can be *external*, such as background noise, a competing voice, or variability in speech signals due to within- and across-speaker variations (Uchanski and Braida, 1998), or internal (e.g.,

neural noise, or fluctuations in attention over time). For tractability, and justified by the central-limit theorem (see: Green and Swets, 1966), the noise is traditionally assumed to have a Gaussian probability distribution.

Depending on whether the noise that limits performance is assumed to occur before or after the combination of information across channels, two models are obtained: the independent-noises model, and the post-combination noise model.

2.4.2.1. The independent-noises model. In this model, d' for the combined-stimulation case (hereafter denoted as d'_C) is related to d' for lowpass-filtered stimulation alone (hereafter denoted as d'_L) and to d' for vocoded stimulation alone (hereafter denoted as d'_V) by,

$$d'_{C_IN} = \sqrt{d'^2_L + d'^2_V}. \quad (2)$$

The subscript, IN, stands for “independent-noises” model. The PCs measured in the L and V conditions can be used to estimate d'_L and d'_V by inverting (numerically) the following integral equation, which describes the relationship between PC and d' for the m -alternative forced-choice (mAFC) identification task (see: Green and Birdsall, 1958; Green and Dai, 1991).

$$PC = \int_{-\infty}^{+\infty} \phi(z - d') \Phi^{m-1}(z) dz. \quad (3)$$

In this equation, $\phi(\cdot)$ denotes the standard normal probability density function, $\Phi^{m-1}(\cdot)$ denotes the cumulative standard normal distribution raised to the power $m - 1$, where m is the number of response alternatives.

Based on a reanalysis of Kryter's (1962) data on speech intelligibility as a function of set-size, Müsch and Buus (2001) found that, when the number of items (e.g., words) in a speech-recognition test is large (e.g., one hundred or more), setting m to 8000 leads to more accurate predictions than setting it to the size of the stimulus set. A similar effect was observed by Green and Birdsall (1958), who suggested that when the stimulus set is large and its contents are not known in advance to the listener, performance is determined, not by the number of stimulus alternatives, but by the size of the listener's active vocabulary. Müsch and Buus (2001) estimated the size of this active vocabulary to be equal to 8000. Accordingly, in the analyses described below, the parameter m was initially set to 8000. However, additional analyses were also performed, in which, either the value of m was treated as a free parameter, or uncertainty regarding the value of m was explicitly taken into account, using a Bayesian approach (described in the Appendix).

Combining Eqs. (2) and (3), the prediction of the independent-noises model for the combined condition is given by,

$$P_{C_IN} = \int_{-\infty}^{+\infty} \phi\left(z - \sqrt{d'^2_L + d'^2_V}\right) \Phi^{m-1}(z) dz. \quad (4)$$

2.4.2.2. The late-noise model. Whereas the independent-noises model assumes that the only significant source of noise occurs *before* information is combined across the L and V signals, the “late-noise” model assumes that the only significant source of performance-limiting noise occurs *after* the combination. In this model, d' for the combined (C) case is predicted using the following equation.

$$d'_{C_LN} = d'_L + d'_V. \quad (5)$$

where the subscript, LN, stands for “late-noise.”

Combining Eqs. (3) and (5), the PC prediction of the late-noise model for the case of combined stimulation is obtained as,

$$P_{C_LN} = \int_{-\infty}^{+\infty} \phi(z - d'_L - d'_V) \Phi^{m-1}(z) dz. \quad (6)$$

3. Results

3.1. PCs measures in human listeners

Fig. 1 provides an overview of the data. The upper panel shows how PC varied as a function of N and SNR, with CF as a parameter, for the three stimulation modes: V, L, and C. Many of the trends that are apparent in this figure are as expected based on earlier findings. In particular PC usually increased with CF, SNR and, for the V and C stimulation modes, it also increased with N . Also consistent with previous findings, PCs measured in the C stimulation mode were generally (with few exceptions) higher than the PC measured (in corresponding SNR, CF, and when applicable, N , conditions) for the L or V stimulation modes.

The statistical significance of these observations was confirmed by the results of repeated-measures ANOVAs, as detailed in the following five sub-sections. These analyses were preceded by Mauchly's test of sphericity. Whenever the sphericity assumption was violated, the Greenhouse-Geisser correction was applied; the F and p values reported in this article include this correction, as needed. Although the results reported below were obtained using analyses performed directly on the PC data, parallel analyses were performed on d' values computed using Eq. (3) (with m set to 8000). Unless noted otherwise, these analyses yielded qualitatively similar outcomes and, in the interest of space, only the results of analyses performed on PC data are reported.

3.1.1. Analysis of results for the L stimulation mode

A two-way (SNR \times CF) ANOVA on the PCs corresponding to the L stimulation mode showed significant main effects of SNR ($F(3, 69) = 154.90, p < 0.0005$) and CF ($F(3, 69) = 84.69, p < 0.0005$), and a significant interaction ($F(9, 207) = 4.50, p < 0.0005$).

3.1.2. Analysis of results for the V stimulation mode

A three-way ($N \times \text{SNR} \times \text{CF}$) ANOVA on the PCs corresponding to the V stimulation mode showed significant main effects of N ($F(3, 69) = 246.87, p < 0.0005$), SNR ($F(3, 69) = 83.08, p < 0.0005$), and CF ($F(2, 46) = 7.39, p = 0.002$), as well as significant interactions between N and SNR ($F(9, 207) = 60.40, p < 0.0005$), SNR and CF ($F(6, 138) = 8.83, p < 0.0005$), and a three-way interaction ($F(18, 414) = 2.71, p = 0.019$). The N -by-CF interaction failed to reach statistical significance ($F(6, 138) = 2.43, p = 0.057$).

3.1.3. Analysis of results for the C stimulation mode

A three-way ($N \times \text{SNR} \times \text{CF}$) ANOVA on the PCs corresponding to the C stimulation mode showed significant main effects of N ($F(3, 69) = 76.60, p < 0.0005$), SNR ($F(3, 69) = 83.08, p < 0.0005$), and CF ($F(2, 46) = 148.27, p < 0.0005$), as well as significant interactions between N and SNR ($F(9, 207) = 3.41, p = 0.004$), N and CF ($F(6, 138) = 12.67, p < 0.0005$), SNR and CF ($F(6, 138) = 3.60, p = 0.008$), and a three-way interaction ($F(18, 414) = 1.90, p = 0.048$). The same ANOVA, when applied on d' , also showed significant main effects of N ($F(3, 69) = 152.890, p < 0.0005$), SNR ($F(3, 69) = 309.11, p < 0.0005$), and CF ($F(2, 46) = 146.21, p < 0.0005$). However, except for the N -by-CF interaction ($F(6, 138) = 9.05, p < 0.0005$), this analysis showed no significant two- or three-way interaction.

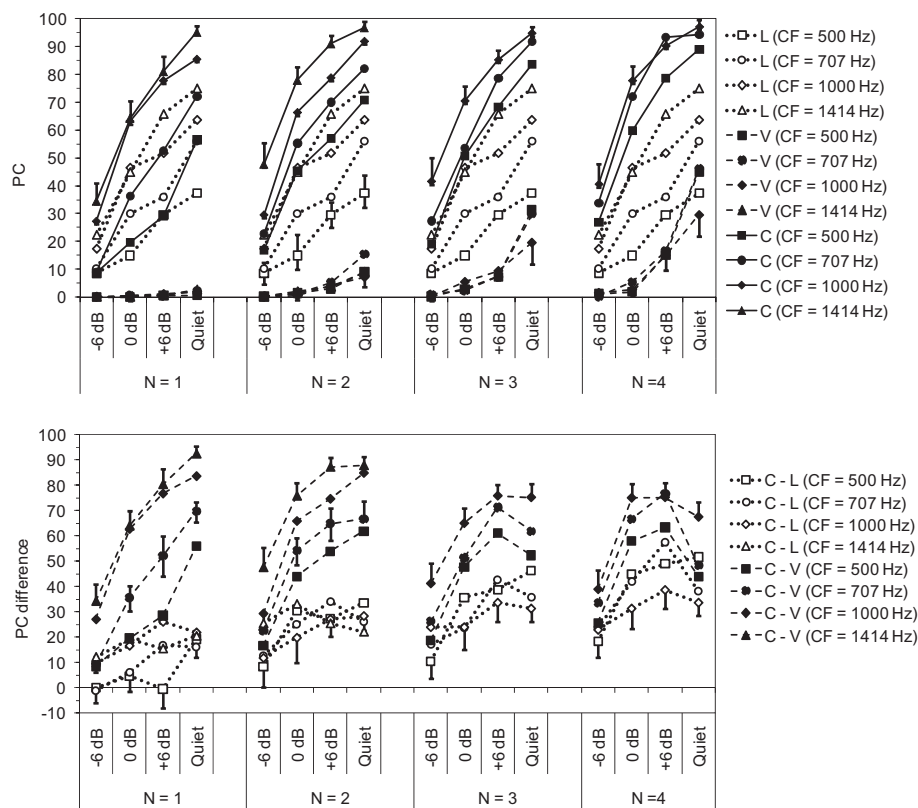


Fig. 1. PC in L, V, and C conditions, and combined-stimulation advantage, as a function of N and SNR, with CF as the parameter. The different CF conditions are illustrated by different symbols, as indicated in the key. The different N and SNR conditions are listed underneath the abscissa. Note that N refers to conditions that involved vocoded stimuli. The data obtained in conditions involving only lowpass-filtered stimuli (empty symbols in the upper panel) were duplicated across N conditions in the upper panel, to facilitate comparisons. The PC differences, which are shown in the lower panel, were computed by subtracting the PCs measured in the L or V conditions from the PCs measured in corresponding C conditions. The error bars show the upper and/or lower bounds of the 95%-confidence intervals (bootstrap) of the mean PC across the 24 listeners for a subset of conditions. These conditions were selected so as to minimize overlap and avoid clutter. They include conditions in which the mean PCs were the lowest or the highest in each N condition; the 500-Hz CF for the $N = 2$ condition (L) in the upper panel; and the 707-Hz CF for $N = 1$ and $N = 2$ condition (C–V) in the lower panel. Note that some error bars for this subset of conditions are not visible, due to their small size relative to the symbols.

3.1.4. Comparisons between L and C stimulation modes

Planned contrasts analyses comparing PC measured in the L stimulation mode against the PCs measured in corresponding SNR and CF conditions of the C stimulation mode, separately for each N , showed significant differences between the two stimulation modes for all N s ($F(1, 23) = 83.04$, $p < 0.0005$ for $N = 1$; $F(1, 23) = 475.36$, $p < 0.0005$ for $N = 2$; $F(1, 23) = 901.00$, $p < 0.0005$ for $N = 3$; and $F(1, 23) = 872.65$, $p < 0.0005$ for $N = 4$).

3.1.5. Comparisons between V and C stimulation modes

A four-way (stimulation mode $\times N \times \text{SNR} \times \text{CF}$) ANOVA on PCs measured using the V and C stimulation modes showed a significant main effect of the stimulation mode ($F(1, 23) = 3529.91$, $p < 0.0005$). Significant interactions were also observed between the stimulation mode and the following factors: N ($F(3, 69) = 24.168$, $p < 0.0005$), SNR ($F(3, 69) = 176.76$, $p < 0.0005$), CF ($F(2, 46) = 148.94$, $p < 0.0005$). Additionally, the stimulation-mode factor was involved in significant three-way interactions with N and SNR ($F(9, 207) = 17.73$, $p < 0.0005$), and N and CF ($F(6, 138) = 7.13$, $p < 0.0005$); the three-way interaction between stimulation mode, SNR, and CF failed to reach statistical significance ($F(6, 138) = 2.33$, $p = 0.057$). Finally, the four-way (stimulation mode $\times N \times \text{SNR} \times \text{CF}$) interaction was significant ($F(18, 414) = 2.07$, $p = 0.035$). A parallel analysis on d' yielded qualitatively identical outcomes, except for the four-way interaction, which was not statistically significant ($F(18, 414) = 1.66$, $p = 0.111$).

3.2. Model predictions

In this section, the predictions of the above-described cue-combination models are compared to the PCs measured for the C stimulation mode. The predicted PCs were obtained by setting the variables P_L and P_V in the equations in Section 2.4 to the mean PCs measured in L and V conditions in the listeners. For the SDT models, predictions could not be computed when the mean PC for the L or V presentation mode was equal to zero, as this led to an undefined d' . To circumvent this problem, mean PC values that were equal to zero were replaced by $1/m$ (i.e., the chance rate) prior to the computation of d' , resulting in a d' of zero for that condition. Only five out of a total of 72 mean PC values had to be replaced in this way. Four of the replaced values corresponded to conditions involving one-band V stimuli; the remaining value corresponded to a condition involving two-band V stimuli.

3.2.1. Probability-summation model

The predictions of the probability-summation model are shown in the upper panel of Fig. 2 (empty symbols), together with the mean PCs measured in the listeners for the C stimulation mode (filled symbols). The differences between the data and the predictions are shown in the lower panel. These differences were obtained by subtracting the predicted PCs from the observed PCs, so that positive values indicate that listeners' performance exceeded the predictions of the model. With a few exceptions, which

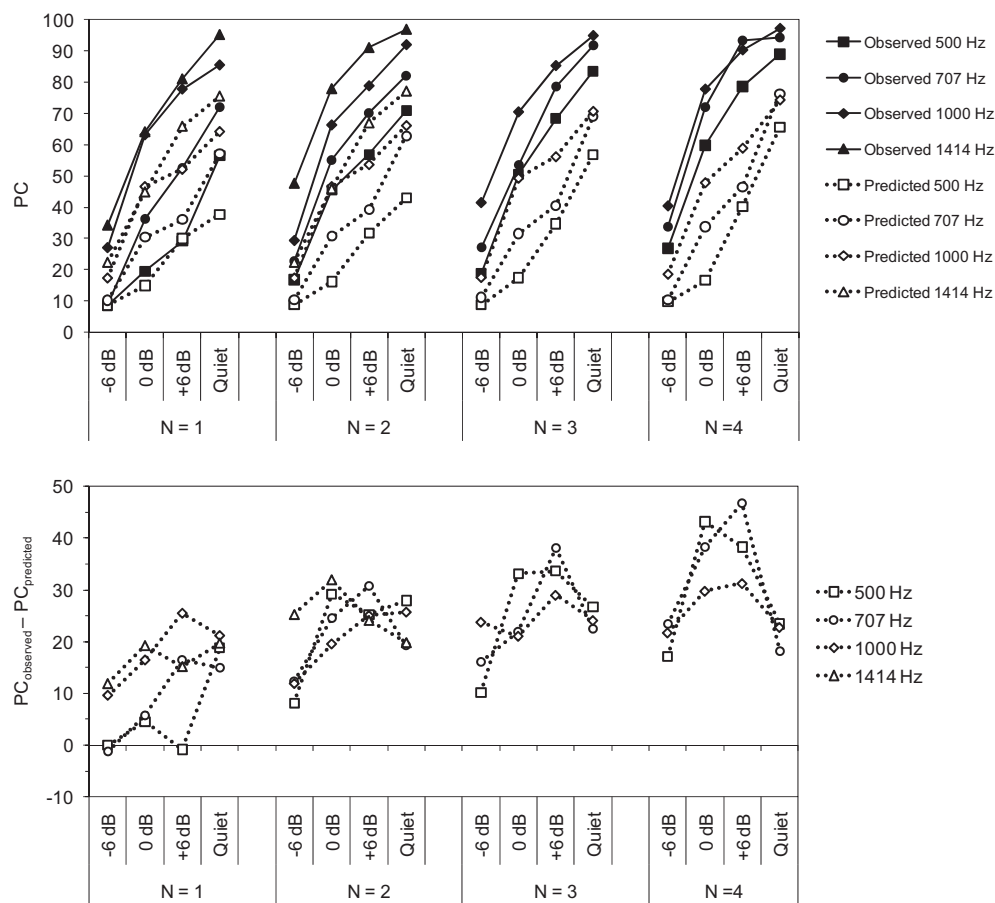


Fig. 2. Comparison between the PCs measured for C stimuli in human listeners, and the PCs predicted by the probability-summation model. Upper panel: Mean PCs of the human listeners for C stimuli (filled symbols, replotted from Fig. 1), and model predictions (empty symbols). Lower panel: Differences between observed and predicted PCs. These differences were obtained by subtracting the predicted PCs from the observed PCs, so that positive values indicate conditions in which listeners' performance was better than predicted by the model, and negative values indicate the opposite effect.

occurred in conditions where N was equal to 1, the mean PCs of the human listeners were higher than predicted by the probability-summation model. In most conditions, the discrepancy between data and predictions was larger than 10 percentage points. In some conditions—especially, when N was equal to 4—the discrepancies were as large as 40–50 percentage points. For this model, the RMS difference between data and predictions across all N , SNR, and CF conditions was equal to 24 percentage points.

3.2.2. Independent-noises model

Fig. 3 shows the predictions of the independent-noises model with the parameter m set to 8000. Although these predictions were almost always lower than the mean PCs of the listeners, the deviations between data and predictions were generally smaller than for the probability-summation model (compare the lower panels of Figs. 2 and 3). The RMS difference between data and predictions was equal to 13 percentage points. This is approximately half the RMS error for the probability-summation model (24 percentage points). However, note that, for many of the conditions tested, the predictions of the independent-noises model were more than 10 percentage points lower than the mean PCs of the human listeners.

As mentioned above, the predictions in Fig. 3 were obtained with the parameter m , set to 8000. As explained in Section 2.2, this choice was based on the conclusion of an earlier study, which estimated the size of listeners' active vocabulary in open-set speech tests at about 8000 words (Müsch and Buus, 2001). However, it is

unclear whether this conclusion applies to the conditions and listeners tested in the current study. To explore the possibility that the independent-noises model could fit the listeners' data more closely for different settings of the parameter m , a gradient-descent algorithm was used to find the value of m that minimized the RMS error between data and predictions. The results, which are shown in Fig. 4, reveal that when m was adjusted in this way, the independent-noises model could fit the human listeners' data with an error of less than 10 percentage points in most of the conditions tested, and an overall RMS error of only 7 percentage points. For comparison, the RMS deviation between the listeners' mean PCs and the upper (or lower) bounds of the 95%-confidence intervals around these mean PCs was equal to 6 percentage points; thus, the RMS error between data and predictions was not appreciably larger than the variability of the data. However, it is worth noting that the m value that was found to yield such good agreement between data and predictions was considerably larger than the estimate obtained by Müsch and Buus (2001), being equal to 870,000. Possible explanations for this result are considered in the discussion.

3.2.3. Late-noise model

The predictions of the late-noise model with the parameter m set to 8000 are shown in Fig. 5. In contrast to the probability-summation and independent-observations models, this model almost invariably over-predicted listeners' performance. The RMS error between data and predictions was equal to 18 percentage points.

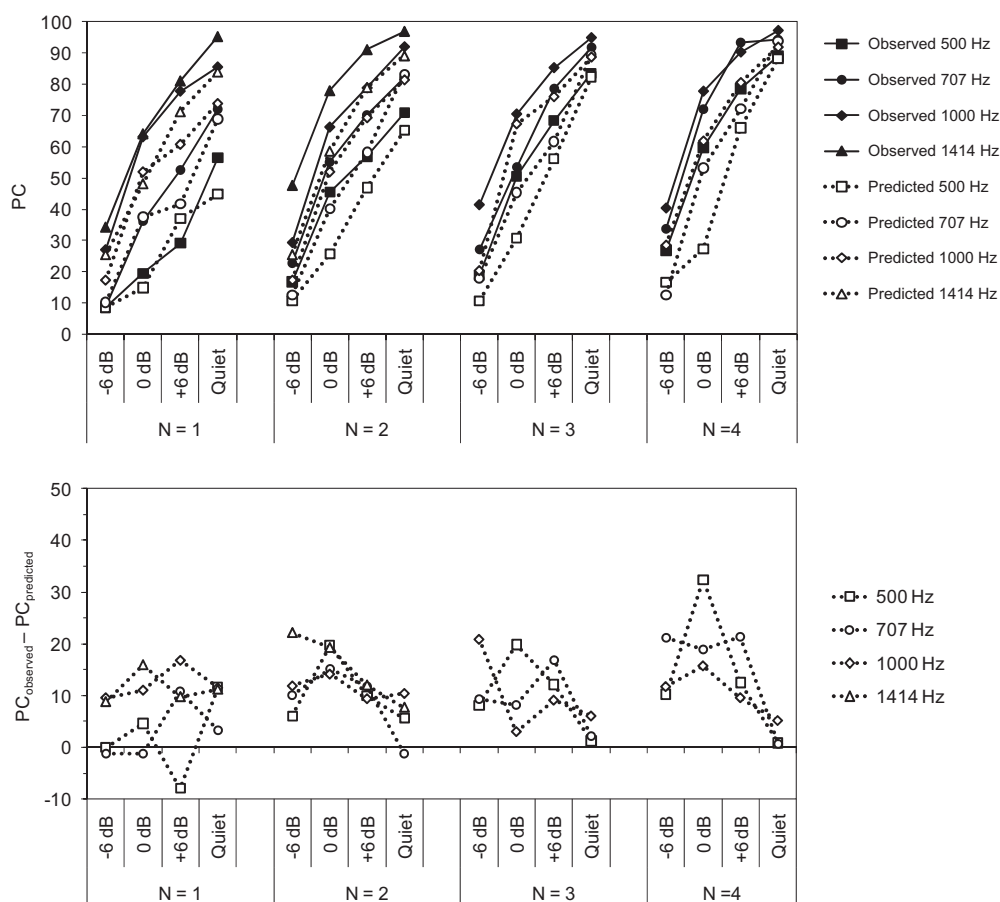


Fig. 3. Comparison between human listeners' performance and the predictions of the independent-noises model, with the value of the parameter, m , set to 8000. The format of this figure is the same as that of Fig. 2.

Fig. 6 shows the predictions that were obtained for the late-noise model when the value of the parameter, m , was adjusted to minimize the RMS error between the data and the predictions. The value of m that was found to minimize the RMS error was equal to 653. For this setting of m , the RMS error between the predictions of the model and the data was equal to 11 percentage points.

3.3. Bayesian model comparison

The finding that smaller RMS errors could be obtained for the two Gaussian-SDT models than for the probability-summation model was not entirely unexpected, since the former two contain one free parameter, whereas the latter contains none. To compare the ability of these models to “explain” the data (in a statistical sense), while taking into account the different numbers of free parameters, we used a Bayesian model-comparison procedure. The mathematical details of this procedure are described in the Appendix. One advantage of the Bayesian approach is that “Occam’s razor” principle, which penalizes models for having more degrees of freedom, is implemented naturally and automatically via the marginalization of the likelihood over the space of model parameters (see: MacKay, 2003, chapter 28, pp. 343–356). Other advantageous features of the Bayesian approach to model-based analyses of psychophysical data are explained in Rouder and Lu (2005). One such advantage, which was especially important in the context of the current study, is that Bayesian inference provides a theoretically principled way to deal with correct response counts equal to zero. This was especially important here, because the predictions of the

models considered in this study are based upon estimates of probabilities of correct response in V conditions, in which correct response counts of zero were sometimes observed. With relatively small numbers of trials per condition, it is possible to obtain correct response counts of zero, even if the underlying probability of a correct response is higher than the rate chance ($1/m$). The Bayesian model-comparison approach described in the Appendix takes this into account.

The outcome of a Bayesian model-comparison analysis is traditionally summarized using Bayes factors. Bayes factors indicate the relative likelihood of a model, compared to another model, given the data. Fig. 7 shows Bayes factors for the independent-noises model, compared to the probability-summation model. The values shown in this figure were obtained by taking the geometric mean, across all listeners, of the “individual” Bayes factors that were computed (using the approach described in the Appendix) for each combination of CF, N , and SNR. Values higher than 1 (indicated by the horizontal line) mean that, for the considered condition, the independent-noises model was more likely than the probability-summation model, given the data; values lower than 1 indicate the opposite. For example, a value of 10 means that, on average across all listeners, the independent-noises model was 10 times more likely to have generated the individual data than the probability-summation model. As can be seen, except for some of the lowest-SNR, lowest-CFs, and lowest- N conditions, the mean Bayes factors in Fig. 7 were generally higher than 1. On average across all conditions and listeners, the mean Bayes factor was 41.73, indicating that the independent-noises model was about

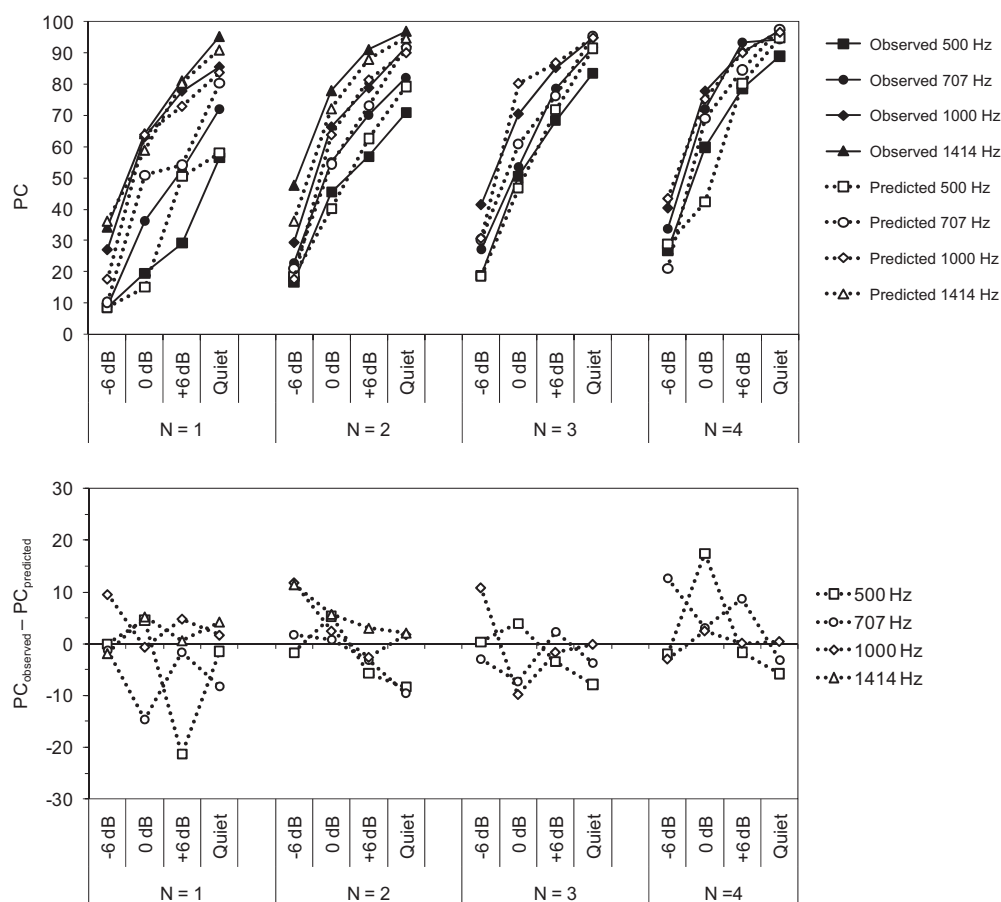


Fig. 4. Comparison between human listeners' performance and the predictions of the independent-noises model, with the value of the parameter, m , adjusted to minimize the mean-squared error between data and predictions. The format of this figure is the same as that of Fig. 2.

42 times more likely to have generated the individual data than the probability-summation model. A similar comparison between the late-noise model and the probability-summation model yielded a mean Bayes factor of 26.08. The mean Bayes factor for the comparison between the independent-noises model and the late-noise models was smaller, being equal to 1.60.

In interpreting these values and those shown in Fig. 7, it is important to note that they are means of Bayes factors computed based on individual data. The "group-level" Bayes factors, which were computed by taking into account the data of all listeners at once (as explained in the Appendix), were several orders of magnitude larger than these "individual" Bayes factors. The geometric mean (across all conditions) of the group-level Bayes factor for the independent-noises model compared to the probability-summation model was equal to $5.85\text{E}+38$. The geometric mean group-level Bayes factor for the late-noise model compared to the probability-summation model was equal to $1.69\text{E}+33$. The geometric mean group-level Bayes factor for the independent-noise model compared to the late-noise model was equal to $3.47\text{E}+05$. According to Jeffreys' scale (Jeffreys, 1961), Bayes factors of 6 or more provide "strong evidence" for one model over another, and Bayes factor larger than 100 provide "conclusive evidence."

4. Discussion

Consistent with previous findings (e.g., Chang et al., 2006; Kong and Carlyon, 2007; Li and Loizou, 2008; Qin and Oxenham, 2006),

the results of this study showed generally higher speech-recognition performance with combined (lowpass-filtered + vocoded) stimulation than with lowpass-filtered or vocoded stimuli alone. This "combined-stimulation advantage" was observed under a wide variety of conditions, including in quiet and in a realistic noise background (cafeteria), for different SNRs (ranging from -6 to $+6$ dB), and different CFs (ranging from 500 to 1414 Hz). The numbers of vocoder bands (N) that were tested in this study were relatively small, ranging from 1 to 4. Interestingly, even with so few vocoder bands, speech-recognition performance was generally higher when both vocoded and lowpass-filtered signals were presented, than when only lowpass-filtered signals were presented. These findings confirm and extend earlier demonstrations of combined-stimulation advantages under simulated-EAS conditions in normal-hearing listeners.

The results of the current study are also consistent with the findings of Kong and Carlyon (2007), who were the first to test the "super-additivity" hypothesis rigorously, using a mathematical model. These authors observed that their listeners' performance in combined-stimulation (lowpass-filtered + vocoded) conditions was higher than predicted by the probability-summation model. A similar outcome was obtained in the current study. However, as noted above, the probability-summation model is largely sub-optimal. The two Gaussian-SDT models that were evaluated in this study both predicted higher performance in combined-stimulation conditions than the probability-summation model, and even when taking into account the fact that these models included a free parameter (while the probability-summation model did not), they

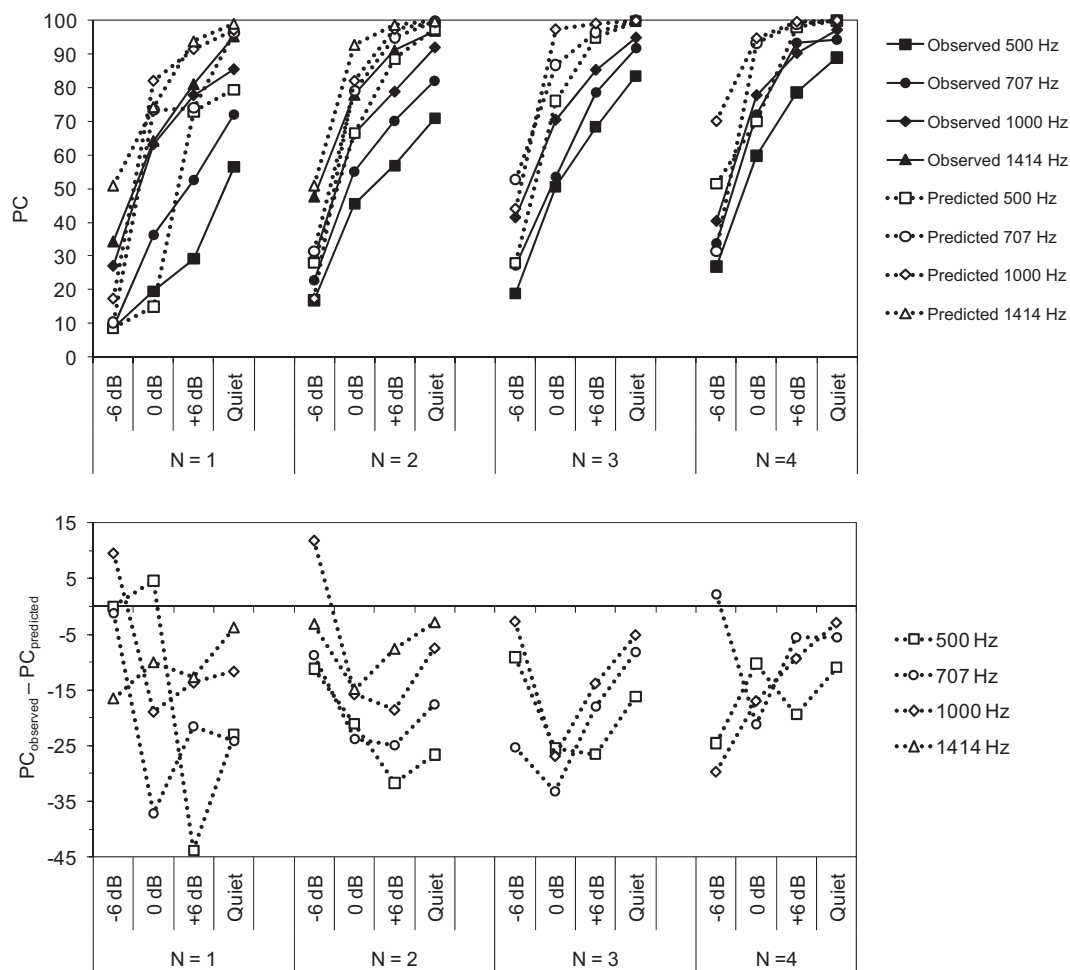


Fig. 5. Comparison between human listeners' performance and the predictions of the late-noise model with the value of the parameter, m , set to 8000. The format of this figure is the same as that of Fig. 2.

were found to provide a far more plausible account of the human listeners' data than the probability-summation model. This outcome suggests that future efforts to develop predictive models of human listeners' speech-recognition performance in combined-stimulation conditions should build upon Gaussian-SDT models, rather than upon the probability-summation model.

When the parameter, m , in the two Gaussian-SDT models was adjusted to minimize the error between the predictions and the data, both models were found to be able to fit the human listeners' speech-recognition performance in combined-stimulation conditions well. The RMS error between data and predictions (across all tested conditions) was equal to 7 percentage points for the independent-noises model, and to 11 percentage points for the late-noise model. Since neither of these models involves any interaction in the combination of information, this indicates that, for the listeners and stimulus conditions tested in this study, the "combined-stimulation advantage" can be explained without invoking "constructive interactions" in the perceptual processing of the low-frequency and vocoded signals. We emphasize that this result does not invalidate explanations of the combined-stimulation advantage that imply a form of constructive interaction. However, it indicates that such explanations may not always be needed. A similar conclusion was reached in a recent study by Kong and Braida (2011). These authors concluded that their listeners' speech-recognition performance in simulated-EAS conditions could be accounted for using "pre-labeling" models, which were based on

Gaussian-SDT assumptions. As those evaluated in the current study, these SDT models did not assume any interaction in the processing of lowpass-filtered and vocoded signals, but they did assume optimal (or quasi-optimal) combination of the cues contained in these signals. More recently, Michey and Oxenham (submitted for publication) found that Gaussian-SDT models could account for the data of Kong and Carlyon (2007). Nonetheless, it is important to acknowledge that the use of a syllable identification task with single-word stimuli, and of a noise vocoder with a 50-Hz cutoff frequency for the envelope lowpass filter (which probably removed even the weakest pitch cues), may have created an experimental environment that was perhaps more conducive to simple additive effects than others. Additional work is needed to further clarify the conditions under which the observed benefits of combined stimulation can, or cannot, be accounted for without positing some form of interaction. Model-based analyses such as those described here and in other articles (e.g., Braida, 1991; Kong and Braida, 2011; Kong and Carlyon, 2007; Michey and Oxenham, submitted for publication; Müsch and Buus, 2001; Ronan et al., 2004), should certainly play a key role in this endeavor.

Two issues with the conclusion that the two Gaussian-SDT models tested in this study could explain the data deserve mention. First, the value of m that was needed to minimize the RMS error between the predictions of the independent-noises model and the listeners' data was quite large, being equal to 870,000. This is more than two orders of magnitude larger than the estimate of

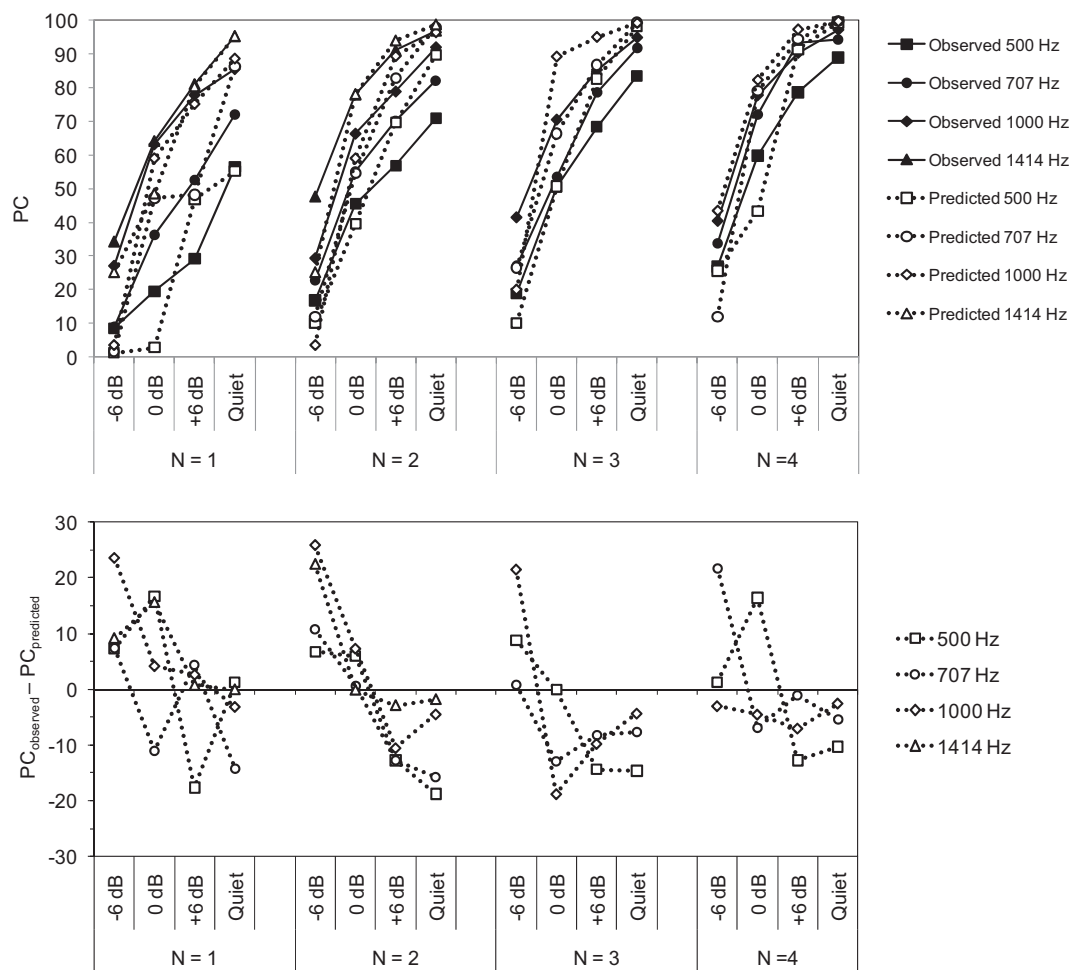


Fig. 6. Comparison between human listeners' performance and the predictions of the late-noise model with the value of the parameter, m , adjusted to minimize the mean-squared error between data and predictions. The format of this figure is the same as that of Fig. 2.

Müsch and Buus (2001), and it appears to be unrealistically large for an estimate of the size of listeners' active vocabulary—although it is important to note that the nature and number of speech “templates” that can be stored in the human brain, and that may be needed to make speech-recognition possible under a wide variety of acoustic

conditions, are currently unknown. This limits the appeal of the independent-noises model. Although the m value needed to minimize the RMS error between data and predictions for the late-noise model was considerably smaller (being equal to 653), the assumption on which this model is based, according to which the

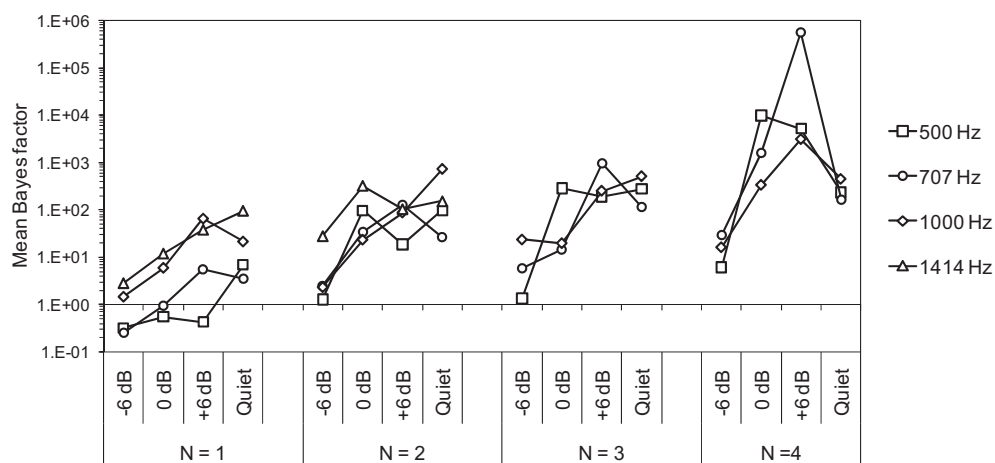


Fig. 7. Geometric mean Bayes factors for the comparison between the independent-noises and probability-summation models. Values higher than 1 (indicated by the horizontal line) indicate that, on average across all listeners, for the considered combination of N , SNR, and CF, the data measured in the experiment were more likely to have been generated by the independent-noises model than by the probability-summation model; values lower than 1 indicate the opposite. Importantly, these Bayes factors refer to *individual* data; the Bayes factors that were computed on the group data (assuming statistical independence across listeners) were orders of magnitude higher.

predominant source of performance-limiting noise occurs after low-frequency and vocoded cues are combined, also seems unrealistic. In low-SNR conditions, where relatively high levels of external noise are present, the noise that limits performance is likely introduced *before*, rather than *after*, information is combined across channels. This limits the appeal of the late-noise model somewhat. However, note that these limitations can be overcome, simply, by combining pre- and post-combination noise sources within the same model, and by letting the magnitude of the pre-combination noise decrease as the SNR increases. A possibility, which we did not consider in the current study, but which it would be interesting to investigate in future experiments, is that increasing external noise increases the uncertainty of the listener, and forces the listener to search through a larger set of templates (which may be thought of as an increase in the listeners' active vocabulary).

Two important limitations of this study must be pointed out. Firstly, we only considered models that did not involve interactions in the perceptual processing of low-frequency and vocoded cues. Our decision on this point was in accord with the stated goal of this study: to test whether the EAS advantage could be accounted for quantitatively without invoking interactions. The decision can be further justified *a posteriori*, given that the answer to this question was found to be positive—thus making the need to search for alternative (and less parsimonious) explanations less pressing. However, in future work, it would be interesting to investigate models of the EAS advantage that involve across-channel interactions. Moreover, it would be important to examine the impact of correlations (e.g., temporal envelope correlations) or, more generally, statistical dependencies, between low-frequency and vocoded (or electric) speech signals. Secondly, and finally, an important caveat, which applies to all simulated-EAS studies, relates to use of noise-vocoding and lowpass-filtering in normal-hearing listeners to simulate CI processing and residual low-frequency hearing in hearing-impaired listeners. Important aspects and consequences of cochlear damages, such as reduced frequency selectivity at low frequencies (Faulkner et al., 1990; Moore, 1985; Moore et al., 1997; Tyler, 1986) and injury-induced neural degeneration (Kujawa and Liberman, 2009), which can greatly limit the benefit of combined acoustic and electric stimulation in hearing-impaired individuals, are not taken into account in these simulations. Therefore, while EAS simulations in normal-hearing listeners provide a test-bed for the development and the evaluation of explanatory or predictive models of the EAS advantage, studies in CI listeners are crucially needed to validate these models.

Acknowledgments

This work was supported by Vibrant Med-El France (Doctoral Research Grant CIFRE 266/2007 to F.S.), the French National Center for Scientific Research (CNRS), NIH R01 DC05216 (author C.M.) and "Laboratoire Audition Conseil" (22 rue Constantine, Lyon 69001, France) Dr. Christopher Brown and an anonymous reviewer provided useful comments on an earlier version of the manuscript.

Appendix. Bayesian model comparison

In Bayesian model comparison, the posterior probabilities of two models, M_A and M_B , given the data, D , are compared by forming the ratio (Gelman et al., 1995; Jaynes, 2003),

$$\frac{P(M_A|D)}{P(M_B|D)} = \frac{P(D|M_A)P(M_A)}{P(D|M_B)P(M_B)} \quad (A1)$$

where $P(D|M_A)$ and $P(D|M_B)$ are the conditional probabilities of the data given the model, and $P(M_A)$ and $P(M_B)$ are the model prior probabilities. When none of the models is favored *a priori* (as was

the case in the current application), the prior probabilities are equal, and the ratio of posterior probabilities, $P(M_A|D)/P(M_B|D)$, equals the ratio of likelihoods, $P(D|M_A)/P(D|M_B)$. The latter ratio, known as the Bayes factor, is computed by integrating over model parameters, for each model:

$$\frac{P(D|M_A)}{P(D|M_B)} = \frac{\int_{\theta_A} P(D|\theta_A, M_A)P(\theta_A|M_A)d\theta_A}{\int_{\theta_B} P(D|\theta_B, M_B)P(\theta_B|M_B)d\theta_B} \quad (A2)$$

In the current application of this framework, the data were numbers of correct responses for lowpass, vocoded, and combined (lowpass + vocoded) stimuli. These numbers of correct responses are denoted as n_L^{ij} , n_V^{ij} , and n_C^{ij} (where the superscripts, i and j , index the listener and the condition), and are assumed to be drawn from binomial distributions with parameters, n , p_L^{ij} , p_V^{ij} , and p_C^{ij} . For clarity, and using probability-theory notation, we have:

$$n_L^{ij}|p_L^{ij}, n \sim \text{binomial}(p_L^{ij}, n), \quad (A3)$$

$$n_V^{ij}|p_V^{ij}, n \sim \text{binomial}(p_V^{ij}, n), \quad (A4)$$

and

$$n_C^{ij}|p_C^{ij}, n \sim \text{binomial}(p_C^{ij}, n). \quad (A5)$$

The shared parameter, n , corresponds to the number of trials per condition per listener; this number is fixed and known ($n = 20$). The parameters, p_L^{ij} , p_V^{ij} , and p_C^{ij} , are the underlying (or "true") probabilities of a correct response. These probabilities are "latent" variables; they are not observed. Although their most-likely (or maximum-*a-posteriori*) values can be inferred based on the measured numbers of correct responses, for the purpose of evaluating the right-hand side of Eq. (A2), these variables are just "nuisance parameters," which must be integrated over to compute the conditional probabilities of interest.

The different models described in Section 2.4 define specific relationships between p_L^{ij} , p_V^{ij} , and p_C^{ij} . Specifically, for the probability-summation model,

$$p_C^{ij} = 1 - (1 - p_L^{ij})(1 - p_V^{ij}). \quad (A6)$$

For the independent-noises model,

$$p_C^{ij} = f_m^{-1}\left(\sqrt{f_m^2(p_L^{ij}) + f_m^2(p_V^{ij})}\right). \quad (A7)$$

where $f_m(\cdot)$ denotes the transformation from d' to PC for the m AFC task, as defined by the integral Eq. (3), and $f_m^{-1}(\cdot)$ denotes the inverse of this transformation. Finally, for the late-noise model,

$$p_C^{ij} = f_m^{-1}\left(f_m(p_L^{ij}) + f_m(p_V^{ij})\right). \quad (A8)$$

In general terms, we can write,

$$p_C^{ij} = g_k(p_L^{ij}, p_V^{ij}). \quad (A9)$$

where the subscript, k , refers to the considered model (1: probability-summation model; 2: independent-noises model; 3: late-noise model).

Together with Eq. (A9), and with the additional assumption that p_L^{ij} and p_V^{ij} are marginally independent, the conditional-dependence

statements (A3)–(A5) define a directed acyclic graph (Bishop, 2006; Jordan, 1999). By examining the structure of this graph, it can be determined that, for the probability-summation model,

$$P(N_L^{ij}, N_V^{ij}, N_C^{ij} | M_k) = \int_0^1 B_n(n_L^{ij}, x_L) P(x_L) \int_0^1 B_n(n_V^{ij}, x_V) \times P(x_V) B_n(n_C^{ij}, f_k(x_L, x_V)) dx_L dx_V, \quad k = 1, \quad (A10)$$

where N_L^{ij} , N_V^{ij} , and N_C^{ij} denote the random variables, of which n_L^{ij} , n_V^{ij} , and n_C^{ij} are observed realizations; x_L and x_V are “dummy” integration variables, which correspond to p_L^{ij} and p_V^{ij} ; and $B_n(q, p)$ is the probability distribution function with parameters n (the number of trials) and p (the probability of success) evaluated at q (the number of successes),

$$B_n(q, p) = \frac{n!}{q!(n-q)!} p^q (1-p)^{n-q}. \quad (A11)$$

Uniform prior distributions were placed on parameters p_L^{ij} and p_V^{ij} . Consequently, the terms $P(x_L)$ and $P(x_V)$ in Eq. (A10) were both equal to one, and,

$$P(N_L^{ij}, N_V^{ij}, N_C^{ij} | M_k) = \int_0^1 B_n(n_L^{ij}, x_L) \int_0^1 B_n(n_V^{ij}, x_V) \times B_n(n_C^{ij}, f_k(x_L, x_V)) dx_L dx_V, \quad k = 1. \quad (A12)$$

The independent-noises model and the late-noise model have one more parameter than the probability-summation model. This parameter, m , enters in the transformation between d' and PC for the mAFC task (Eq. (3)). It corresponds to the number of independent templates, against which the listener is comparing incoming speech signals, and can be thought of as the size of the listener’s “active vocabulary” (Green and Birdsall, 1958; Müsch and Buus, 2001). Since m is not known *a priori*, this extra parameter must also be “integrated out” in the calculation of the likelihoods of the independent-noises and late-noise models. Accordingly, for these models, the likelihood was computed as follows.

$$P(N_L^{ij}, N_V^{ij}, N_C^{ij} | M_k) = \sum_0^\infty P(m) \int_0^1 B_n(n_L^{ij}, x_L) \int_0^1 B_n(n_V^{ij}, x_V) \times B_n(n_C^{ij}, f_k(x_L, x_V)) dx_L dx_V dm, \quad k = \{2, 3\}. \quad (A13)$$

The infinite upper limit in the sum on the right-hand side of Eq. (A13) reflects the fact that m can theoretically be any non-negative integer. However, the size of listeners’ active vocabulary is most certainly finite. For practical purposes, uncertainty concerning this parameter was modeled using a uniform-discrete prior probability distribution on m , such that values of m equally spaced on an octave scale between 125 and 125×2^{13} were regarded as equally likely *a priori*.

The integrals in Eqs. (A12) and (A13) were replaced by sums, and evaluated numerically. The Bayes factors were then obtained simply by taking ratios of the conditional probabilities, $P(N_L^{ij}, N_V^{ij}, N_C^{ij} | M_k)$, corresponding to two different models (e.g., the probability-summation model and the independent-noises model). This was done for all pairs of models. Importantly, these Bayes factors were computed separately for each listener and condition. The mean Bayes factors reported in the Results section were computed as the geometric mean, across all listeners, of these individual Bayes factors.

In addition to “individual” Bayes factors, “group-level Bayes factors” were computed as

$$\frac{P(D|M_A)}{P(D|M_B)} = \frac{\prod_{j=1}^{n_{\text{subs}}} P(N_L^{ij}, N_V^{ij}, N_C^{ij} | M_A)}{\prod_{j=1}^{n_{\text{subs}}} P(N_L^{ij}, N_V^{ij}, N_C^{ij} | M_B)}, \quad (A14)$$

where n_{subs} is the number of subjects, and M_A and M_B refer to the two models being compared.

References

- Başkent, D., Chatterjee, M., 2010. Recognition of temporally interrupted and spectrally degraded sentences with additional unprocessed low-frequency speech. *Hear. Res.* 270, 127–133.
- Bishop, C., 2006. *Pattern Recognition and Machine Learning*. Springer, Berlin.
- Boothroyd, A., Nittrouer, S., 1988. Mathematical treatment of context effects in phoneme and word recognition. *J. Acoust. Soc. Am.* 84, 101–114.
- Braida, L.D., 1991. Crossmodal integration in the identification of consonant segments. *Q. J. Exp. Psychol. A* 43, 647–677.
- Brokx, J.P., Nooteboom, S.G., 1982. Intonation and the perceptual separation of simultaneous voices. *J. Phonetics* 10, 23–36.
- Brown, C.A., Bacon, S.P., 2009a. Low-frequency speech cues and simulated electric-acoustic hearing. *J. Acoust. Soc. Am.* 125, 1658–1665.
- Brown, C.A., Bacon, S.P., 2009b. Achieving electric-acoustic benefit with a modulated tone. *Ear Hear.* 30, 489–493.
- Brown, C.A., Bacon, S.P., 2010. Fundamental frequency and speech intelligibility in background noise. *Hear. Res.* 266, 52–59.
- Büchner, A., Schüssler, M., Battmer, R.D., Stover, T., Lesinski-Schiedat, A., Lenarz, T., 2009. Impact of low-frequency hearing. *Audiol. Neurotol.* 14 (Suppl. 1), 8–13.
- Carlyon, R.P., 1996. Encoding the fundamental frequency of a complex tone in the presence of a spectrally overlapping masker. *J. Acoust. Soc. Am.* 99, 517–524.
- Carroll, J., Zeng, F.G., 2007. Fundamental frequency discrimination and speech perception in noise in cochlear implant simulations. *Hear. Res.* 231, 42–53.
- Chang, J.E., Bai, J.Y., Zeng, F.G., 2006. Unintelligible low-frequency sound enhances simulated cochlear-implant speech recognition in noise. *IEEE Trans. Biomed. Eng.* 53, 2598–2601.
- Chen, F., Loizou, P.C., 2010. Contribution of consonant landmarks to speech recognition in simulated acoustic-electric hearing. *Ear Hear.* 31, 259–267.
- Ching, T.Y., Incerti, P., Hill, M., 2004. Binaural benefits for adults who use hearing aids and cochlear implants in opposite ears. *Ear Hear.* 25, 9–21.
- Culling, J.F., Darwin, C.J., 1993. Perceptual separation of simultaneous vowels: within and across-formant grouping by F0. *J. Acoust. Soc. Am.* 93, 3454–3467.
- Cullington, H.E., Zeng, F.G., 2011. Comparison of bimodal and bilateral cochlear implant users on speech recognition with competing talker, music perception, affective prosody discrimination, and talker identification. *Ear Hear.* 32, 16–30.
- Dorman, M.F., Spahr, A.J., Loizou, P.C., Dana, C.J., Schmidt, J.S., 2005. Acoustic simulations of combined electric and acoustic hearing (EAS). *Ear Hear.* 26, 371–380.
- Faulkner, A., Rosen, S., Moore, B.C., 1990. Residual frequency selectivity in the profoundly hearing-impaired listener. *Br. J. Audiol.* 24, 381–392.
- Fletcher, H., 1953. *Speech and Hearing in Communication*. Krieger, Huntington, NY.
- Fournier, J.E., 1951. *Audiométrie vocale*. Maloine, Paris.
- Gantz, B.J., Turner, C.W., 2003. Combining acoustic and electrical hearing. *Laryngoscope* 113, 1726–1730.
- Gantz, B.J., Turner, C.W., 2004. Combining acoustic and electrical speech processing: lowa/Nucleus hybrid implant. *Acta Otolaryngol.* 124, 344–347.
- Gantz, B.J., Turner, C.W., Gfeller, K., 2004. Expanding cochlear implant technology: combined electrical and acoustical speech processing. *Cochlear Implants Int.* 5 (Suppl. 1), 8–14.
- Gantz, B.J., Turner, C.W., Gfeller, K.E., 2006. Acoustic plus electric speech processing: preliminary results of a multicenter clinical trial of the lowa/Nucleus Hybrid implant. *Audiol. Neurotol.* 11 (Suppl. 1), 63–68.
- Gelman, A., Carlin, J., Stern, H., Rubin, D., 1995. *Bayesian Data Analysis*. Chapman and Hall/CRC, Boca Raton, Florida.
- Gfeller, K.E., Olszewski, C., Turner, C., Gantz, B., Oleson, J., 2006. Music perception with cochlear implants and residual hearing. *Audiol. Neurotol.* 11 (Suppl. 1), 12–15.
- Green, D.M., Birdsall, T.G., 1958. The Effect of Vocabulary Size on Articulation Score. Technical Memorandum No. 81 and Technical Note AFCRC-TR-57-58. University of Michigan: Electronic Defense Group.
- Green, D.M., Dai, H.P., 1991. Probability of being correct with 1 of M orthogonal signals. *Percept. Psychophys.* 49, 100–101.
- Green, D.M., Swets, J.A., 1966. *Signal Detection Theory and Psychophysics*. Krieger, New York.
- von Ilberg, C., Kiefer, J., Tillein, J., Pfenningdorff, T., Hartmann, R., Stürzebecher, E., Klinke, R., 1999. Electric-acoustic stimulation of the auditory system. New technology for severe hearing loss. *J. Otorhinolaryngol. Relat. Spec.* 61, 334–340.
- Jaynes, E.T., 2003. *Probability Theory: The Logic of Science*. Cambridge University Press, New York.
- Jeffreys, H., 1961. *The Theory of Probability*, third ed. Clarendon Press, Oxford.

- Jordan, M.I., 1999. *Learning in Graphical Models*. MIT Press, Cambridge, MA.
- Kong, Y.Y., Braida, L.D., 2011. Cross-frequency integration for consonant and vowel identification in bimodal hearing. *J. Speech Lang. Hear. Res.* 54, 959–980.
- Kong, Y.Y., Carlyon, R.P., 2007. Improved speech recognition in noise in simulated binaurally combined acoustic and electric stimulation. *J. Acoust. Soc. Am.* 121, 3717–3727.
- Kong, Y.Y., Stickney, G.S., Zeng, F.G., 2005. Speech and melody recognition in binaurally combined acoustic and electric hearing. *J. Acoust. Soc. Am.* 117, 1351–1361.
- Kryter, K.D., 1962. Methods for the calculation and use of the articulation index. *J. Acoust. Soc. Am.* 34, 1689–1697.
- Kujawa, S.G., Liberman, M.C., 2009. Adding insult to injury: cochlear nerve degeneration after “temporary” noise-induced hearing loss. *J. Neurosci.* 29, 14077–14085.
- Li, N., Loizou, P.C., 2008. A glimpsing account for the benefit of simulated combined acoustic and electric hearing. *J. Acoust. Soc. Am.* 123, 2287–2294.
- MacKay, D.J.C., 2003. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, Cambridge, UK.
- Macmillan, N.A., Creelman, C.D., 2005. *Detection Theory: A User's Guide*, second ed. Erlbaum, Mahwah, NJ.
- Michéyl, C., Oxenham, A.J., 2009. Revisiting evidence for “super-additive” benefits of combined stimulation, submitted for publication.
- Mok, M., Grayden, D., Dowell, R.C., Lawrence, D., 2006. Speech perception for adults who use hearing aids in conjunction with cochlear implants in opposite ears. *J. Speech Lang. Hear. Res.* 49, 338–351.
- Moore, B.C., Vickers, D.A., Glasberg, B.R., Baer, T., 1997. Comparison of real and simulated hearing impairment in subjects with unilateral and bilateral cochlear hearing loss. *Br. J. Audiol.* 31, 227–245.
- Moore, B.C., 1985. Frequency selectivity and temporal resolution in normal and hearing-impaired listeners. *Br. J. Audiol.* 19, 189–201.
- Müsch, H., Buus, S., 2001. Using statistical decision theory to predict speech intelligibility. I. Model structure. *J. Acoust. Soc. Am.* 109, 2896–2909.
- Nittrouer, S., Boothroyd, A., 1990. Context effects in phoneme and word recognition by young children and older adults. *J. Acoust. Soc. Am.* 87, 2705–2715.
- Pelli, D.G., 1985. Uncertainty explains many aspects of visual contrast detection and discrimination. *J. Opt. Soc. Am. A* 2, 1508–1532.
- Pirenne, M.H., 1943. Binocular and uniocular threshold of vision. *Nature* 152, 698–699.
- Qin, M.K., Oxenham, A.J., 2006. Effects of introducing unprocessed low-frequency information on the reception of envelope-vocoder processed speech. *J. Acoust. Soc. Am.* 119, 2417–2426.
- Ronan, D., Dix, A.K., Shah, P., Braida, L.D., 2004. Integration across frequency bands for consonant identification. *J. Acoust. Soc. Am.* 116, 1749–1762.
- Rouder, J.N., Lu, J., 2005. An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychon. Bull. Rev.* 12, 573–604.
- Treisman, M., 1998. Combining information: probability summation and probability averaging in detection and discrimination. *Psychol. Methods* 3, 252–265.
- Turner, C.W., Gantz, B.J., Vidal, C., Behrens, A., Henry, B.A., 2004. Speech recognition in noise for cochlear implant listeners: benefits of residual acoustic hearing. *J. Acoust. Soc. Am.* 115, 1729–1735.
- Tyler, R.S., 1986. Frequency resolution in hearing-impaired listeners. In: Moore, B.C.J. (Ed.), *Frequency Selectivity in Hearing*. Academic Press, London, pp. 309–371.
- Uchanski, R.M., Braida, L.D., 1998. Effects of token variability on our ability to distinguish between vowels. *Percept. Psychophys.* 60, 533–543.
- Wickens, T., 2001. *Elementary Signal Detection Theory*. Oxford University Press, Oxford.

II.7. Evaluation de différentes stratégies de codage de la parole par des simulations d'implant EAS

(Article soumis à Hear Res. : Seldran F, Gallego S, Thai-Van H, Berger-Vachon C. A test of different speech-coding strategies in simulated electric-acoustic hearing)

Un problème émergent avec la stimulation électro-acoustique est l'augmentation des interactions qui résultent d'un espace inter-électrodes réduit sur les porte-électrodes conçus pour l'EAS. Pour remédier à cela, il pourrait être intéressant d'utiliser une stratégie de codage qui vise à limiter les interactions, à savoir une stratégie à extraction de maxima ou une stratégie à cadence de stimulation réduite.

Pour répondre à cette question, nous avons réalisé à l'aide d'un vocoder hybride, une étude sur 30 sujets normo-entendants qui écoutaient des sons filtrés passe-bas, vocodés, et combinés (vocodés + passe-bas) dans le silence et dans le bruit à des RSB allant de -6 à +12 dB. Nous avons élaboré notre vocoder de sorte à évaluer l'effet du nombre de maxima (variant entre 2/10 et 10/10). Nous avons également utilisé différents chevauchements entre nos différentes trames d'analyses pour simuler différentes cadences de stimulation par l'implant (entre 250 rafraîchissements par seconde soit un rafraîchissement toutes les 4 ms, et 1000 rafraîchissements par seconde soit un rafraîchissement toutes les 1 ms).

Nos résultats montrent que l'augmentation du nombre de maxima augmente significativement les performances, en particulier en présence de bruit. L'augmentation de la cadence de rafraîchissement produit un effet modéré et pas toujours significatif sur les performances. En supposant que les résultats obtenus par simulation donnent une indication de ce qui se produit chez les implantés utilisateurs de l'EAS, nous pouvons supposer que si un sujet peut exploiter son audition résiduelle, une stratégie de codage avec un nombre réduit d'informations ne dégrade pas les performances.

Ces résultats suggèrent que pour l'EAS, l'utilisation d'une stratégie à extraction de pics utilisée avec une cadence réduite permettrait de limiter les interactions.

A test of different speech-coding strategies in simulated electric-acoustic hearing

Fabien Seldran,^{1,2,3,4*} Stéphane Gallego,^{1,2,3} Hung Thai-Van,^{1,2,3,5} Christian Berger-Vachon^{1,2,3}

1. INSERM U1028, Lyon Neuroscience Research Center, PACS Team (Speech, Audiology, Communication Health), Lyon, F-69000, France

2. CNRS UMR5292, Lyon Neuroscience Research Center, PACS Team (Speech, Audiology, Communication Health), Lyon, F-69000, France

3. University Lyon 1, Lyon, F-69000, France

4. Vibrant Med-El Hearing Technology GmbH, 400 Ave. Roumanille, BP 309, Sophia-Antipolis 06906, France

5. Audiology and ENT Department, Edouard Herriot Hospital, Lyon F-69437, France

*Corresponding author: INSERM U1028, CNRS UMR5292, Centre de Recherche en Neurosciences de Lyon, Equipe Audition, Hôpital Edouard Herriot, Pavillon U, Place d'Arsonval, F-69437 Lyon Cedex 03, France, Tel: +33.4.72.11.05.03, Fax: +33.4.72.11.05.04

email addresses: fseldran@yahoo.fr (F. Seldran), sgallego@hotmail.fr (S. Gallego), hthaivan@gmail.com (H. Thai-Van), christian.berger-vachon@univ-lyon1.fr (C. Berger-Vachon)

Short title: Coding strategies in EAS simulation

Abbreviations

ACE: advance combination encoder

CI: cochlear implant

CIS: continuous interleaved sampling

EAS: electro-acoustic stimulation

FFT: fast Fourier transform

FPS: frames per second

RMS: root-mean-square

SPEAK: spectral peak

SNR: signal-to-noise ratio

Abstract

Advances in speech signal processing for cochlear implants (CIs) have led to the development of “N of M” coding strategies, whereby only a subset of electrode channels are stimulated during a given time frame. The two main parameters for this type of strategy are the number of active channels (N) and the frame rate; the latter usually corresponds to the per-channel stimulation rate. The goal of this study was to investigate how these two parameters influence speech-recognition performance in the situation where the listener has access to unprocessed low-frequency components of the speech signal—as happens with combined electric-plus-acoustic stimulation (EAS). To this aim, a simulation approach was used, wherein normal-hearing (NH) subjects were presented with highpass-filtered (> 650 Hz) speech signals processed using a dynamic vocoder to simulate N-of-M coding and/or with the lowpass-filtered (< 650 Hz) but otherwise unprocessed components of the same speech signal. Recognition performance for such simulated combined, lowpass-only, and highpass-vocoded-only signals was measured in quiet and in the presence of a realistic noise background (cafeteria noise), for SNRs ranging from -6 to $+12$ dB, in 30 listeners, N values ranging from 2 to 10 (out of 10), and frame rates ranging from 250 frames per second (FPS) to 1000 FPS. The results showed that increasing N from 2 to 10 channel could produce large beneficial effects on speech-recognition performance, especially in the presence of background noise, both for simulated E-only conditions and for simulated EAS conditions. Increasing frame rate from 250 FPS to 1000 FPS produced relatively small, and not always statistically significant, improvements in performance. To the extent that results obtained using acoustic simulations of EAS are indicative of the performance that can be achieved by some patients using a CI in combination with a hearing-aid, these findings suggest that when low-frequency acoustic information is available, high levels of speech-recognition performance

can be attained using a small number of channels ($N = 2$ out of 10) and a slow frame rate (250 FPS).

Keywords: electro-acoustic stimulation; cochlear implant; speech-coding strategies; N-of-M

1. Introduction

Owing to advances in surgical techniques, it is now possible to preserve low-frequency acoustic hearing in some cochlear-implant (CI) patients (e.g., Fraysse *et al.*, 2006; Hodges *et al.*, 1997). This has led to the development of a new approach for the treatment of partial deafness, which is referred to as “electric-acoustic stimulation” (EAS) (e.g., Büchner *et al.*, 2009; Gantz & Turner, 2003; Gfeller *et al.*, 2006; Von Ilberg *et al.*, 1999). In this approach, a short electrode array is inserted into the basal region of the cochlea (which normally responds to high-frequency sounds) while attempting to avoid damage to more apical regions of the cochlea (which respond to lower-frequency sounds), so that the latter can still be stimulated acoustically.

The development of EAS had led CI manufacturers to design shorter electrode arrays. For example, the Iowa/Nucleus Hybrid CI System features a 10- or 6-mm electrode array (Gantz & Turner, 2003). Med-El’s Flex^{EAS} features a 20.9-mm array—compared to 26.4 mm for the standard (non-EAS) array (Adunka *et al.*, 2004; Hochmair *et al.*, 2006). One limitation of short electrode arrays, however, is that they provide a smaller number of electrode channels than longer arrays or, that the inter-electrode spacing is shorter. For instance, the Iowa/Nucleus Hybrid 10-mm CI System features six channels (compared to between 12 and 22 for standard arrays), meaning that the maximum inter-electrode spacing is about 1.6 mm (Gantz *et al.*, 2005). Med-El’s Flex^{EAS} provides 12 stimulation channels, with an inter-electrode spacing of 1.9 mm (compared to 2.4 mm for Med-El’s standard array). A smaller inter-electrode spacing is likely to result in increased interactions between the electrode channels. Channel interactions are one of the most significant factors limiting speech intelligibility in CI listeners (e.g., Arnolder *et al.*, 2007; Bierer & Middlebrooks, 2004; Favre & Pelizzone, 1993; Fu & Shannon, 1999; Lim *et al.*, 1989; Shannon, 1983; Stickney *et al.*, 2006; Tong & Clark, 1986; White *et al.*, 1984), and may

explain why, in some patients, speech-recognition saturates, and even decreases, as the number of active electrode channels increases beyond about 4-8 (Berger-Vachon et al., 1992; Brill et al., 1997).

A commonly used approach to limit channel interactions in CIs involves using non-simultaneous pulses. One of the earliest examples of sound-coding strategies of this type is the “continuous interleaved sampling” (CIS) strategy (Wilson *et al.*, 1991). In this strategy, electrodes are stimulated serially, avoiding temporal overlap. More recent advances in non-simultaneous stimulation have led to the development of so called “N of M” strategies, such as the “spectral maxima sound processor” (SMSP; McDermott *et al.*, 1992), “spectral peak picking” (SPEAK; Seligman & McDermott, 1995; Skinner *et al.*, 1994;), or “advanced combination encoder” (ACE; Skinner *et al.*, 2002). The basic principle of “N of M” strategies is that only a subset (N out of M) of electrode channels are stimulated during a given time frame (Wilson *et al.*, 1988). In almost all cases, the N electrode channels that are selectively stimulated correspond to the most-energetic spectral bands, or “spectral maxima.” Several studies have found that N-of-M strategies can yield higher speech-recognition performance than CIS (e.g., Buechner *et al.*, 2009; McDermott *et al.*, 1992; McKay & McDermott, 1993; Skinner *et al.*, 1994; Whitford *et al.*, 1995). Since N of M strategies apparently provide an efficient way of limiting channel interactions without sacrificing intelligibility, they may prove especially useful for EAS. However, to our knowledge, the potential benefits of using an N-of-M strategy with a relatively small number of available channels (e.g., $M = 10$), and in the presence of unprocessed low-frequency acoustic information, has not been systematically investigated.

Another important parameter, which can also influence across-channel interactions, is the stimulation rate. Higher stimulation rates usually promote channel interactions (De Balthasar *et*

al., 2003; Middlebrooks, 2004; Shannon, 1990; Tykocinski *et al.*, 1995). However, it has been suggested that higher stimulation rates can also have advantageous consequences, including a better representation of the temporal fine structure, and of the fastest amplitude fluctuations in the temporal envelope, of speech signals (Rubinstein *et al.*, 1999; Vandali *et al.*, 2000; Wilson *et al.*, 1991). In N-of-M strategies, the per-channel stimulation rate usually coincides with the frame rate.

The overarching goal of this study was to investigate the influence of the two main parameters of an N-of-M coding strategy, N and frame rate, on speech-recognition performance in a simulated-EAS context. As in previous simulated-EAS studies, CI processing was simulated in normal-hearing (NH) listeners using a vocoder, while residual low-frequency hearing was simulated using lowpass filtering (e.g., Başkent & Chatterjee, 2010; Brown & Bacon, 2009; Chang *et al.*, 2006; Chen & Loizou, 2010; Dorman *et al.*, 2005; Gantz & Turner, 2003; Kong & Carlyon, 2007; Li & Loizou, 2008; Qin & Oxenham, 2006; Seldran *et al.*, in press; Turner *et al.*, 2004). One important difference between these earlier simulated-EAS studies and the current one is that, here, rather than using fixed channels, channels were dynamically updated to simulate an N-of-M strategy, as done by, e.g., Loizou *et al.* (2000a) and Dorman *et al.* (2000; 2002). The influence of N and frame rate on speech-recognition performance was evaluated, both, in quiet and in the presence of a realistic noise background (cafeteria noise). To determine whether the influence of N and frame rate on speech-recognition performance was fundamentally different depending on whether or not low-frequency information was provided, simulated E-only conditions were also tested, in addition to simulated EAS conditions.

2. Methods

2.1 Participants

A total of 30 subjects (14 females, 16 males) aged 19 to 29 years (mean = 23.3 years) took part in the study. Fifteen subjects (7 females, 8 males) aged 19 to 28 years (mean = 22.6) participated in Experiment 1. Fifteen other subjects (7 females, 8 males) aged 19 to 29 years (mean = 23.9 years) participated in Experiment 2. All had normal-hearing. Normal hearing was defined as hearing thresholds better than 20 dB HL at octave frequencies between 250 and 8000 Hz. Testing took place in a double-wall sound-attenuating booth. In accordance with the Declaration of Helsinki, written informed consent was obtained from the subjects prior to their inclusion into the study. The study was approved by the local Ethics Committee (Hospices Civils de Lyon, France, Number ID RCB: 2008-A01479-46).

2.2 Stimuli

The stimuli were disyllabic French words uttered by a single male speaker (Fournier, 1951), and recordings of “cafeteria noise,” which contained a mixture of chatter with background noises produced by cutlery, plates, and glasses, typical of a crowded restaurant or cafeteria. With the exception of the “quiet” testing conditions, in which the words were presented without any background noise, the word and noise signals were mixed prior to processing. In Experiment 1, the following three signal-to-noise ratios (SNRs) were tested: -6, 0, and +6 dB. In Experiment 2, the SNRs were: 0, +6, and +12 dB. In addition, both experiments included a “quiet” condition. For data-analysis and data-exposition purposes, the quiet condition was treated as one of the SNR conditions. The stimuli were presented at 65 dB SPL.

2.3 Signal processing for Experiment 1

All signal processing was implemented in Matlab[®] (MathWorks, Natick, MA). For processing, the signals were down-sampled to 16.7 kHz, yielding a Nyquist frequency of 8.3 kHz. This frequency corresponds approximately to the upper limit of the acoustic frequency range typically used in CI processors. Low-frequency acoustic stimulation was simulated by lowpass filtering the signals using a 6th-order Butterworth filter with a 3-dB cutoff frequency of 650 Hz. CI processing was simulated as follows. First, the signal was split into overlapping “frames” of 128 points each. With the 16.7-kHz sampling frequency, this corresponded to a frame duration of approximately 7.7 ms (as in Cochlear’s SPEAK strategy). To simulate different encoding rates, the percentage of overlap between consecutive frames was varied systematically. Depending on the condition being tested, it was equal to 50% (corresponding to an inter-frame interval of approximately 4 ms, which corresponds to a frame rate of 250 FPS, as in Cochlear’s SPEAK strategy), 75% (corresponding to an inter-frame interval of approximately 2 ms, which corresponds to a frame rate of 500 FPS), or 87.5% (corresponding to an inter-frame interval of approximately 1 ms, which corresponds to a frame rate of 1000 FPS).

The signals contained in each frame were multiplied by Hanning windows, and the resulting windowed signals were subjected to Fast-Fourier transform (FFT). FFT bins corresponding to frequencies comprised between 650 and 8300 Hz were grouped into 10 frequency bands, corresponding to 10 simulated electrode channels.¹ The center frequencies of the 10 bands were equidistant on the bark scale. The relationship between Hz and barks is given by the following equation (Tjornov, 1971).

$$z = 6.7 \times \arg \sinh[(f - 20)/600], \quad (1)$$

where z is in barks, and f is in Hz. This relationship corresponds approximately to that between frequency and cochlear place (Traunmüller, 1990). According to the above formula, 650 Hz

corresponds to 6.1 barks, and 8333 Hz corresponds to 22.3 barks, so that each of the 10 frequency bands was 1.6-bark wide. The frequencies of the bands are listed in Table 1.

Table 1

Then, for each frequency band, the root-mean-square (RMS) amplitude of the FFT components contained within that band was computed. This operation concluded the “analysis” part of the processing chain. For the “synthesis” part, the RMS amplitudes corresponding to the N most energetic bands for the current time frame were used to scale 128-point Hanning windows corresponding to the times frames that were used to reconstruct the temporal envelopes corresponding to the different bands. For this experiment, N values of 2, 4, 6, and 10 (out of 10) were tested. For the 10- N least energetic bands, all elements of the corresponding 128-point vectors were set to 0. The “synthesis frames” corresponding to a given frequency band were then delayed and added according to an overlap-add procedure (Allen, 1977). The resulting temporal-envelope signals were then lowpass-filtered using a 6th-order Butterworth filter with a 3-dB cutoff frequency of 150 Hz. Finally, the filtered temporal envelopes were used to modulate carrier signals. For bands 3 to 10, the carrier signals were bandpass noises obtained by filtering broadband Gaussian noise into frequency bands (the “synthesis bands”), which had the same center frequencies as the corresponding analysis bands, but which were 150-Hz narrower (i.e., 75 Hz on each side) than the corresponding analysis bands, to avoid interactions between modulation side-bands (Caroll & Zeng, 2007). Because bands 1 and 2 were less than 150-Hz wide, for these two bands, pure-tones with frequencies equal to the center frequency of the corresponding analysis band were used as carriers, instead of noise bands. Each vocoded sub-band signal was scaled by a factor equal to the ratio of the RMS of the analysis-band signal to

the RMS of the synthesis-band signal in the $N = 10$ condition. Note that the RMS of the synthesis-band signal was less than that of the analysis-band signal for N smaller than 10.

Finally, the amplitude-modulated noise bands (for bands 3-10) and pure tones (for bands 1-2) were summed, and added to the lowpass-filtered signal (below 650 Hz). The signals were resampled at 44.1 kHz before burning to CD.

2.4 Signal processing for Experiment 2

The signal-processing chain used for Experiment 2 was identical to that used for Experiment 1, except for the following. First, a pre-emphasis filter (first-order highpass Butterworth filter with a cutoff frequency of 1200 Hz) was applied to the signals before the extraction of spectral maxima. Secondly, the temporal envelopes that were extracted from the analysis bands were lowpass filtered below 125 Hz, instead of 150 Hz in Experiment 1. Thirdly, to produce the synthesis bands, the lowpass-filtered envelopes were multiplied (each separately) by broadband noise, and then, the resulting amplitude-modulated broadband noises were filtered using sixth-order Butterworth filters, the 3-dB cutoff frequencies of which were equal to the corner frequencies of the corresponding analysis bands; this procedure obviated the need to “trim” the synthesis bands to avoid spectral interactions caused by modulation sidebands (since these sidebands were removed by the post-modulation filtering). Thirdly, in this experiment, only the two extreme (i.e., lowest and highest) frame rates (i.e., 250 FPS and 1000 FPS) and N values (i.e., 2 and 10) were tested. Finally, and importantly, while in Experiment 1 speech recognition performance was measured only for simulated *combined* stimulations (with the lowpass-filtered and vocoded stimuli always presented simultaneously to the listener), in Experiment 2, each listener was tested under three simulated stimulation modes: lowpass only

(simulating acoustic stimulation only); vocoded only (simulating electric stimulation alone); and lowpass+vocoded (simulating EAS stimulation). For brevity, these three stimulation modes are hereafter referred to as A (for “acoustic”), E (for “electric”), and AE (for “acoustic plus electric”).

2.5 Procedure

The words were arranged into 40 lists of 10 words each. In Experiment 1, each subject was presented with one list of 10 words in each test condition. The tests took approximately 45 minutes per listener. The listeners did not receive training prior to the tests. In Experiment 2, each subject was presented with two lists of words per test condition. Subjects were instructed to repeat each word following its presentation. The tests took approximately 90 minutes per listener and the subjects were allowed to take 10-15 minutes breaks after 45 minutes of testing. The listeners received 10 minutes training prior to the tests.

2.6 Apparatus

The signals were recorded on a CD (44.1 kHz sampling frequency, 16-bit quantization). They were presented using a PHILIPS CD723 CD player connected to a MADSEN Orbiter 922 audiometer, which was used to adjust the stimulus presentation level. The stimuli were presented to the subject’s right ear via the right earpiece of TDH39 earphones. The use of monaural rather than binaural stimulation was based on the consideration that most CI listeners are implanted unilaterally.

2.7 Data analysis

Scoring was based upon the number of correctly repeated syllables. For data-visualization purposes, the number of correctly repeated syllables were converted to percent-correct scores, as this is the metric most commonly used to quantify speech-recognition performance. For statistical-analysis purposes, the percent-correct scores were transformed into rationalized-arcsine units (RAU) (Studebaker, 1985) and the resulting RAU scores were analyzed using repeated-measure ANOVAs. The SNR, the stimulation rate, and the number of extracted maxima were used as factors. The ANOVAs were preceded by Mauchly's sphericity test and the Greenhouse-Geisser correction was applied whenever the sphericity assumption was not met.

3. Results

3.1 Experiment 1

The results of this experiment are shown in Figs. 1 and 2. Figure 1 illustrates the effect of parameter N separately for the three frame rates (panels A-C), and with the data averaged across rates (panel D). A three-way (N x frame rate x SNR) repeated-measure ANOVA on the RAU-transformed score showed a significant main effect of N [$F(3, 42) = 21.65, p < 0.0005$], a significant main effect of SNR [$F(3, 42) = 390.96, p < 0.0005$], and a significant interaction between these two factors [$F(9, 126) = 3.06, p = 0.003$]. Looking at Fig. 1, it can be seen that this interaction stemmed from the effect of N being generally larger at the 0-dB SNR, where performance was usually intermediate (50%). For the Quiet condition, PC was at, or close to, ceiling. However, after RAU transformation, a significant effect of N was found even for the Quiet condition [$F(3, 42) = 4.57, p = 0.007$]. It worth noting that performance in Quiet was still at, or close to, ceiling even for N = 2.

Figure 1

No significant main effect of, or interaction with, frame rate was found. This lack of effect of the frame rate is most apparent when looking at Fig. 2. In this figure, the data have been replotted so that the different curves within each panel correspond to different frame rates, while each different panel corresponds to a given N (panels A-D), or to the average across all four N conditions (panel E).

Figure 2

3.2 Experiment 2

The results of this experiment are illustrated in Fig. 3. The two upper panels (panels A and B) make it easiest to see the effect of the frame rate and the differences in performance between the different stimulation modes (A, E, and AE) for each N separately. Note that since the parameter, N, had no bearing on the A stimulation mode, the data for this stimulation mode, which are shown in these two panels, are the same; these data are plotted in both panels to facilitate comparisons of performance between the A stimulation mode and the other modes. The A and AE data are replotted in Panel C to facilitate the visual inspection of the effect of N on performance in the AE condition.

These data were analyzed using a four way (mode x N x rate x SNR) repeated-measure ANOVA. The data for the A stimulation mode could not be included into this analysis, as this would have resulted in an unbalanced design (due to the rate and N factors not being applicable to this stimulation mode); however, additional analyses were performed to compare performance for this stimulation mode with the performance obtained using the other two stimulation modes. As before, the RAU-transformed scores were used as dependent variable. The results of this global analysis showed significant main effects of stimulation mode [$F(1, 14) = 750.92, p <$

0.0005], N [$F(1, 14) = 230.66, p < 0.0005$], frame rate [$F(1, 14) = 12.15, p = 0.004$], and SNR [$F(3, 42) = 591.56, p < 0.0005$]. In addition, significant interactions between stimulation mode and N [$F(1, 14) = 21.60, p < 0.0005$], SNR and N [$F(3, 42) = 15.88, p < 0.0005$], and stimulation mode and SNR [$F(3, 42) = 62.55, p < 0.0005$] were present. No other significant interaction was found.

Figure 3

Given the significant main effect of stimulation mode, the data were further analyzed using three-way (N x rate x SNR) repeated-measures ANOVAs on the data obtained in the E and AE stimulation modes taken separately. The analysis on the data obtained in the E mode showed significant main effects of N [$F(1, 14) = 162.13, p < 0.0005$], rate [$F(1, 14) = 1984.21, p < 0.0005$], and SNR [$F(3, 42) = 414.65, p < 0.0005$]. The SNR x N interaction just failed to reach statistical significance [$F(3, 42) = 616.41, p = 0.0052$]. No other significant interaction was observed for this stimulation mode. For the AE stimulation mode, significant main effects of N [$F(1, 14) = 133.85, p < 0.0005$], rate [$F(1, 14) = 6.90, p = 0.020$], and SNR [$F(3, 42) = 292.23, p < 0.0005$] were present, and the SNR x N interaction was significant [$F(3, 42) = 21.60, p < 0.0005$]. No other significant interactions were found.

To gain further insight into these effects, separate two-way (rate x SNR) repeated-measure ANOVAs were performed on the data corresponding to each N condition for a given stimulation mode (E and, separately, AE). The results of these analyses showed significant main effects of rate for both stimulation modes for N = 2 [for E: $F(1, 14) = 7.50, p = 0.016$; for AE: $F(1, 14) = 6.49, p = 0.023$], and no significant main effect of rate, regardless of the stimulation mode, for N = 10. No significant rate x SNR interactions were observed.

4. Discussion

Considering that the influence of the two parameters, N (number of maxima) and frame rate, on speech-recognition performance for N-of-M coding strategies is an important question for clinicians, manufacturers, and researchers, one expects to find a large number of publications on this topic. Certainly, numerous studies have examined how speech-recognition performance is influenced by the stimulation rate, both in CI patients and in normal-hearing (NH) subjects listening to vocoded stimuli (simulating CI processing) (e.g., Balkany *et al.*, 2007; Buechner *et al.*, 2006; Friesen *et al.*, 2005; Fu and Shannon, 2000; Kiefer *et al.*, 2000; Loizou *et al.*, 2000b; Schatzer *et al.*, 2010; Vandali *et al.*, 2000; Weber *et al.*, 2007; Wilson *et al.*, 2000a, 2000b). Fewer studies have been devoted to investigating how the choice of N , in N-of-M strategies, influences speech recognition performance (e.g., Arora *et al.*, 2011; Buechner *et al.*, 2009; Buechner *et al.*, 2010; Friesen *et al.*, 2005; Shannon *et al.*, 2011). However, these studies almost invariably involved *fixed*, rather than dynamically updated, channels. By contrast, very few studies have been devoted specifically to characterizing the influence of N and/or frame rate on speech recognition for N-of-M strategies (e.g., Buechner *et al.*, 2009; Fishman *et al.*, 1997). Moreover, in general, these two parameters have not been manipulated independently. Lastly, the current study is the first, to our knowledge, to investigate specifically the effects of N and frame rate on speech-recognition performance in simulated EAS conditions—in addition to simulated E-only and A-only conditions. To facilitate reading, the discussion of the results of this study is organized into subsections.

4.1 Effect of N on speech-recognition performance in simulated EAS

The results of the two experiments of this study both showed significant improvements in speech-recognition performance as the parameter N (number of maxima) was increased from 2 to 10 (out of 10 channels). These beneficial effects were observed, both, in simulated-EAS conditions (Experiments 1 and 2), and in simulated E-only conditions (Experiment 2). The effects of N on PC were generally larger in the presence of background noise than for Quiet conditions. For the latter, no (or weak) effects of N were observed, which can be explained simply by the fact that the PCs measured in Quiet conditions were already at, or close to, ceiling, even for $N = 2$ (the smallest N value tested in this study). At 0-dB SNR, the effect of N was often substantial, with PC increasing from about 40-45% on average for $N = 2$, to about 75-80% for $N = 10$; this corresponds approximately to a 40-percentage-point increase. This suggests that, to the extent that the effects of N observed in these two simulation experiments are indicative of the effects of N on speech-recognition performance in CI patients stimulated using EAS, the use of a small N may be sufficient in quiet listening conditions, but inadequate in noisy conditions. This finding is consistent with the results of previous studies, which have shown that while CI patients and NH subjects listening to vocoder simulations can achieve asymptotically high levels of speech recognition in quiet using a small number of fixed channels (e.g., Shannon *et al.*, 1995), in the presence of background noise, a larger number of channels is needed (e.g., Dorman *et al.*, 1997; Loizou *et al.*, 1999). However, importantly, the results of the current study show that, while with $N = 2$ performance in E conditions was close to zero at 0-dB SNR, in AE conditions, the mean PC for the same N and SNR conditions was considerably higher, being comprised between 40% and 50%. It is worth pointing out that percent-correct scores for words in sentences may be higher than the percent-correct scores for isolated disyllabic words measured in this study, due to semantic context effects and increased predictability (see, e.g., Miller and Heise,

1951). Thus, to the extent that these simulations results are indicative of the performance of EAS patients, they suggest that for patients in whom channel interactions are forcing the clinician to use a very small N, it may still be possible to achieve relatively high levels of speech understanding, owing to the provision of low-frequency unprocessed information. However, this conclusion is subject to the following two caveats. First, this study used a relatively high cutoff frequency for the lowpass-filtered acoustic stimulation. Currently, the majority of patients who are considered as candidates for a short-electrode CI have an audiogram cutoff even lower than this. Second, even if advances in surgical techniques make it possible in the near future, to implant with short-electrode arrays patients in whom low-frequency residual hearing extends up to a higher cutoff (e.g., 700-1000 Hz), reduced frequency selectivity or other suprathreshold deficits may still limit the ability of these patients to take advantage of low-frequency information, especially in the presence of background noise.

4.2 Effect of frame rate

Another important question pertaining to the use of N-of-M strategies, besides the issue of choosing an optimal N, is how to select the frame rate. If the frame rate is too slow relative to the rate of spectral change in the signal (e.g., speech), important cues for speech understanding (e.g., formant transitions, or plosive bursts) may be lost. Cochlear's SPEAK strategy, one of the earliest examples of N-of-M strategy, uses a fixed frame rate of 250 frames per second—corresponding to an inter-frame rate of 4 ms, with approximately 50% overlap between consecutive frames, which are approximately 8 ms in duration each. In recent years, there has been a push for using higher stimulation rates, which, for N-of-M strategies, usually implies using a higher frame rate. However, the evidence that high frame or stimulation rates enhance

speech-recognition performance is not entirely clear, and while some studies have found beneficial effects of increasing rate (Buechner *et al.*, 2006; Fu and Shannon, 2000 ; Kiefer *et al.*, 2000; Loizou *et al.*, 2000b; Nie *et al.*, 2006; Wilson *et al.*, 2000a, 2000b), others have not (Balkany *et al.*, 2007; Friesen *et al.*, 2005; Schatzer *et al.*, 2010; Vandali *et al.*, 2000; Weber *et al.*, 2007).

The results of this study are similarly ambiguous with respect to the effects of frame rate. No statistically significant effect of frame rate was observed in Experiment 1, and statistically significant but relatively small effects of frame rate were observed in Experiment 2. The different outcome of the two experiments is likely due to methodological differences. In particular, Experiment 2 involved twice as many words per condition as Experiment 1. The larger number of observations per condition per listener presumably yielded more accurate and less variable estimates of PC, which may have allowed the relatively small influence of frame rate on performance to be revealed in Experiment 2. In addition, differences in signal processing between the two experiments may have influenced the results. Specifically, the use of pre-emphasis filter in Experiment 2, which was done to better simulate CI processing, may have resulted in higher-frequency channels being more likely to be selected (due to their having more energy) in this experiment than in Experiment 1; since the detection of spectral maxima and the selection of channels was performed dynamically, the difference in the relative balance of low- and high-frequency maxima between the two experiments could interact with the effect frame rate.

The finding that increasing frame rate from 250 to 1000 FPS had a relatively small impact on speech-recognition performance in Experiment 2, and no significant effect in Experiment 1, is not entirely surprising, given that the temporal envelopes of the subband signals

were filtered below 125 Hz. This filtering may have limited the benefits of increasing the frame rate. To test this explanation, we performed additional measures, in which we increased the cutoff frequency of the envelope filter to 500 Hz while using the rate of 1000 FPS. The results of these additional measures (not shown here), which were obtained in the 15 listeners of Experiment 2, did not show a significant improvement over the 125-Hz cutoff, 1000-frames-per-second condition. This negative outcome suggests that lowpass filtering the envelope below 125 Hz was not a significant limiting factor of the effect of frame rate on speech-recognition performance in Experiment 2.

An alternative explanation for the finding that increasing the frame rate from 250 to 1000 FPS had either no, or a small, effect on speech-recognition performance is that the lowest frame rate that was used in this study, 250 frames per second—which corresponds to the frame rate used in the SPEAK strategy (Seligman & McDermott, 1995; Skinner *et al.*, 1994)—was already higher than the upper cutoff of the range of modulation frequencies that has been found to be most important for speech intelligibility, which extends from about 0 to 16 Hz, with marginal contributions from higher envelope frequencies (e.g., Drullman *et al.*, 1994; Shannon *et al.*, 1995). From this point of view, the results of the current study, showing no or relatively small effects of frame rate, are consistent with the view that providing information corresponding to high envelope frequencies (> 250 Hz) provides little or no benefit for speech-recognition performance.

4.3 EAS advantage

The results of this study are consistent with those of earlier studies, which have found higher speech-recognition performance in actual or simulated EAS conditions than for electric

(or vocoded) stimulation alone or acoustic (or lowpass-filtered) stimulation alone (e.g., Başkent & Chatterjee, 2010; Brown & Bacon, 2009; Chang *et al.*, 2006; Chen & Loizou, 2010; Dorman *et al.*, 2005; Gantz & Turner, 2003; Kong & Carlyon, 2007; Li & Loizou, 2008; Qin & Oxenham, 2006; Seldran *et al.*, in press; Turner *et al.*, 2004). Looking at Figure 3, it can be seen that performance was generally higher for the AE presentation mode than for the two other modes, indicating an advantage of combined stimulation over non-combined stimulation. Importantly, this “combined stimulation advantage” was observed even for $N = 2$ —suggesting that the use of an N-of-M coding strategy for EAS may be beneficial, even if N is very small. For $N = 2$, the largest benefits of adding A stimulation to E stimulation were observed at intermediate SNRs (+6 and +12 dB), with little benefit (5-10 percentage points) at the smallest SNR tested (0 dB), where performance in E conditions was at, or close, to zero. By contrast, for $N = 10$, the largest “EAS advantage” occurred for the 0-dB SNR, where the A and E stimuli each yielded between 20 and 40% correct when presented on their own.

Since the use of short electrode arrays for EAS places strong constraints on the number of available electrode channels, it seems especially important, and timely, to examine how speech-recognition performance in EAS depends on N and frame rate in an N-of-M coding strategy. Although the current study is a first step toward this important goal, simulations studies alone are not sufficient. Therefore, it will be especially important, in future work, to study the effect of N and frame rate on the speech-recognition performance of patients equipped with a short-electrode CI and a hearing aid.

4.4 Implications for N-of-M parameter setting in EAS

Before considering the potential implications of the results obtained in this study for the choice of N-of-M parameters in CI patients, it is important to acknowledge that the results obtained using simulations of CI processing and EAS such as those used in this study may not generalize to CI patients. Due to differences in auditory-nerve-fiber responses to electric and acoustic stimuli, and other factors (e.g., limited neural survival in CI patients) the same choice of parameters may lead to very different results in CI patients and NH subjects listening to acoustic simulations of CI processing. As Dorman *et al.* (2000) eloquently put it, “The performance of the normal-hearing listeners establishes a benchmark for how well implanted patients could perform if electrode arrays were able to reproduce, by artificial electrical stimulation, the stimulation produced by auditory stimulation of the cochlea, and if patients possessed neural structures capable of responding to the electrical stimulation.” With this caveat in mind, it is worth asking what the findings of the current study may suggest concerning the selection of N and frame rate in EAS patients. We believe that the answer to this question can be summarized as follows.

First, the relatively high levels of speech-recognition performance in quiet with an N of only 2 suggest that the performance of CI listeners with such a small N is not limited upwards by insufficient spectro-temporal information. Therefore, for patients in whom channel interactions represent a major issue, it may be worthwhile trying to set N to a small value. However, it is important to note that, in the presence of background noise, speech-recognition performance dependence markedly on N, and was often substantially lower for $N = 2$ than for $N = 10$. Therefore, in noisy situations, a strategy that uses all available channels (such as CIS) may be preferable than a strategy that activates only a small subset of channels per frame (such as ACE).

Second, the fact that varying the frame rate from 250 to 1000 FPS had either no, or a relatively small, impact on speech-recognition performance in quiet and in noise, and regardless

of whether unprocessed low-frequency information was available or not, suggests that there may be little advantage in using high frame rates in EAS patients (Wilson, 2011). However, it is important to note that this conclusion is based on results obtained using acoustic stimuli, the temporal envelopes of which were lowpass-filtered below 125 Hz. The possibility cannot be ruled out that more marked effects of the frame rate would have been observed if a higher envelope cutoff frequency had been used.

Acknowledgements

This work was supported by Vibrant Med-El France (Doctoral Research Grant CIFRE 266/2007 to FS) and the French National Center for Scientific Research (CNRS). The authors would like to express their gratitude to M. Agrawal for her help with Matlab programming and data collection, to C. Micheyl for his assistance with statistical analyses and writing, and to M. Beliaeff, S. Labassi, B. Philibert, and E. Truy for their support throughout the study.

Footnotes

1. By way of comparison, for the Med-El Flex^{EAS} array, 10 functional electrode channels correspond approximately to an insertion depth of 18 – 22 mm.

References

- Adunka, O., Kiefer, J., Unkelbach, M.H., Lehnert, T., Gstoettner, W. 2004. Development and evaluation of an improved cochlear implant electrode design for electric acoustic stimulation. *Laryngoscope* 114, 1237-41.
- Allen, J.B. 1977. Short Term Spectral Analysis, Synthesis, and Modification by Discrete Fourier Transform. *IEEE Trans on Acoust Speech and Sig Proc ASSP-25*(3), 235-238.
- Arnoldner, C., Riss, D., Baumgartner, W.D., Kaider, A., Hamzavi, J.S. 2007. Cochlear implant channel separation and its influence on speech perception--implications for a new electrode design. *Audiol Neurotol* 12, 313-24.
- Arora, K., Vandali, A., Dowell, R., Dawson, P. 2011. Effects of stimulation rate on modulation detection and speech recognition by cochlear implant users. *Int J Audiol* 50, 123-32.
- Balkany, T., Hodges, A., Menapace, C., Hazard, L., Driscoll, C., Gantz, B., Kelsall, D., Luxford, W., McMenomy, S., Neely, J.G., Peters, B., Pillsbury, H., Roberson, J., Schramm, D., Telian, S., Waltzman, S., Westerberg, B., Payne, S. 2007. Nucleus Freedom North American clinical trial. *Otolaryngol Head Neck Surg* 136, 757-62.
- Başkent, D., Chatterjee, M. 2010. Recognition of temporally interrupted and spectrally degraded sentences with additional unprocessed low-frequency speech. *Hear Res.* 270, 127-133.
- Berger-Vachon, C., Collet, L., Djedou, B., Morgon, A. 1992. Model for understanding the influence of some parameters in cochlear implantation. *Ann Otol Rhinol Laryngol* 101, 42-45.
- Bierer, J.A., Middlebrooks, J.C. 2004. Cortical responses to cochlear implant stimulation: channel interactions. *J Assoc Res Otolaryngol* 5, 32-48.

- Brill, S.M., Gstöttner, W., Helms, J., von Ilberg, C., Baumgartner, W., Müller, J., Kiefer, J. 1997. Optimization of channel number and stimulation rate for the fast continuous interleaved sampling strategy in the COMBI 40+. *Am J Otol* 18 Suppl 6, S104-6.
- Brown, C.A., Bacon, S.P. 2009. Low-frequency speech cues and simulated electric-acoustic hearing. *J Acoust Soc Am* 125, 1658-65.
- Büchner, A., Schüssler, M., Battmer, R.D., Stöver, T., Lesinski-Schiedat, A., Lenarz, T. 2009. Impact of low-frequency hearing. *Audiol Neurotol* 14 Suppl 1, 8-13.
- Buechner, A., Frohne-Buechner, C., Boyle, P., Battmer, R.D., Lenarz, T. 2009. A high rate n-of-m speech processing strategy for the first generation Clarion cochlear implant. *Int J Audiol* 48, 868-75.
- Buechner, A., Frohne-Buechner, C., Gaertner, L., Lesinski-Schiedat, A., Battmer, R.D., Lenarz, T. 2006. Evaluation of Advanced Bionics high resolution mode. *Int J Audiol* 45, 407-16.
- Buechner, A., Frohne-Büchner, C., Gaertner, L., Stoever, T., Battmer, R.D., Lenarz, T. 2010. The Advanced Bionics High Resolution Mode: stimulation rates up to 5000 pps. *Acta Otolaryngol* 130, 114-23.
- Carroll, J., Zeng, F.G. 2007. Fundamental frequency discrimination and speech perception in noise in cochlear implant simulations. *Hear Res* 231, 42-53.
- Chang, J.E., Bai, J.Y., Zeng, F.G. 2006. Unintelligible low-frequency sound enhances simulated cochlear-implant speech recognition in noise. *IEEE Trans Biomed Eng* 53, 2598-601.
- Chen, F., Loizou, P.C. 2010. Contribution of consonant landmarks to speech recognition in simulated acoustic-electric hearing. *Ear Hear* 31, 259-67.

- De Balthasar, C., Boëx, C., Cosendai, G., Valentini, G., Sigrist, A., Pelizzone, M. 2003. Channel interactions with high-rate biphasic electrical stimulation in cochlear implant subjects. *Hear Res* 182, 77-87.
- Dorman, M.F., Loizou, P.C., Fitzke, J., Tu, Z. 2000. Recognition of monosyllabic words by cochlear implant patients and by normal-hearing subjects listening to words processed through cochlear implant signal processing strategies. *Ann Otol Rhinol Laryngol Suppl.* 185, 64-6.
- Dorman, M.F., Loizou, P.C., Rainey, D. 1997. Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. *J Acoust Soc Am* 102, 2403-11.
- Dorman, M.F., Loizou, P.C., Spahr, A.J., Maloff, E. 2002. A comparison of the speech understanding provided by acoustic models of fixed-channel and channel-picking signal processors for cochlear implants. *J Speech Lang Hear Res* 45, 783-8.
- Dorman, M.F., Spahr, A.J., Loizou, P.C., Dana, C.J., Schmidt, J.S. 2005. Acoustic simulations of combined electric and acoustic hearing (EAS). *Ear Hear* 26, 371-80.
- Drullman, R., Festen, J.M., Plomp, R. 1994. Effect of temporal envelope smearing on speech reception. *J Acoust Soc Am* 95, 1053-64.
- Favre, E., Pelizzone, M. 1993. Channel interactions in patients using the Ineraid multichannel cochlear implant. *Hear Res* 66, 150-6.
- Fishman, K.E., Shannon, R.V., Slattery, W.H. 1997. Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. *J Speech Lang Hear Res* 40, 1201-15.
- Fournier JE. 1951. *Audiométrie vocale*. Maloine, Paris.

- Friesen, L.M., Shannon, R.V., Cruz, R.J. 2005. Effects of stimulation rate on speech recognition with cochlear implants. *Audiol Neurotol* 10, 169-84.
- Fu, Q.J., Shannon, R.V. 1999. Effects of electrode location and spacing on phoneme recognition with the Nucleus-22 cochlear implant. *Ear Hear* 20, 321-31.
- Fu, Q.J., Shannon, R.V. 2000. Effect of stimulation rate on phoneme recognition by nucleus-22 cochlear implant listeners. *J Acoust Soc Am* 107, 589-97.
- Gantz, B.J., Turner, C.W. 2003. Combining acoustic and electrical hearing. *Laryngoscope* 113, 1726-30.
- Gantz, B.J., Turner, C.W., Gfeller, K.E., Lowder, M.W. 2005. Preservation of hearing in cochlear implant surgery: advantages of combined electrical and acoustical speech processing. *Laryngoscope* 115, 796-802.
- Gfeller, K.E., Olszewski, C., Turner, C., Gantz, B., Oleson, J. 2006. Music perception with cochlear implants and residual hearing. *Audiol Neurotol* 11 Suppl 1, 12-5.
- Hochmair, I., Nopp, P., Jolly, C., Schmidt, M., Schösser, H., Garnham, C., Anderson, I. 2006. MED-EL Cochlear implants: state of the art and a glimpse into the future. *Trends Amplif* 10, 201-19.
- Hodges, A.V., Schloffman, J., Balkany, T. 1997. Conservation of residual hearing with cochlear implantation. *Am J Otol* 18, 179-83.
- Kiefer, J., von Ilberg, C., Rupprecht, V., Hubner-Egner, J., Knecht, R. 2000. Optimized speech understanding with the continuous interleaved sampling speech coding strategy in patients with cochlear implants: effect of variations in stimulation rate and number of channels. *Ann Otol Rhinol Laryngol* 109, 1009-20.

- Kong, Y.Y., Carlyon, R.P. 2007. Improved speech recognition in noise in simulated binaurally combined acoustic and electric stimulation. *J Acoust Soc Am* 121, 3717-27.
- Li, N., Loizou, P.C. 2008. A glimpsing account for the benefit of simulated combined acoustic and electric hearing. *J Acoust Soc Am* 123, 2287-94.
- Lim, H.H., Tong, Y.C., Clark, G.M. 1989. Forward masking pattern produced by intracochlear electrical stimulation of one and two electrode pairs in the human cochlea. *J Acoust Soc Am* 86, 971-980.
- Loizou, P.C., Dorman, M., Tu, Z. 1999. On the number of channels needed to understand speech. *J Acoust Soc Am* 106, 2097-103.
- Loizou, P.C., Dorman, M.F., Tu, Z., Fitzke, J. 2000a. Recognition of sentences in noise by normal-hearing listeners using simulations of speak-type cochlear implant signal processors. *Ann Otol Rhinol Laryngol* 185, 67-8.
- Loizou, P.C., Poroy, O., Dorman, M. 2000b. The effect of parametric variations of cochlear implant processors on speech understanding. *J Acoust Soc Am* 108, 790-802.
- McDermott, H.J., McKay, C.M., Vandali, A.E. 1992. A new portable sound processor for the University of Melbourne/Nucleus Limited multielectrode cochlear implant. *J Acoust Soc Am* 91, 3367-71.
- McKay, C.M., McDermott, H.J. 1993. Perceptual performance of subjects with cochlear implants using the Spectral Maxima Sound Processor (SMSP) and the Mini Speech Processor (MSP). *Ear Hear* 14, 350-67.
- Middlebrooks, J.C. 2004. Effects of cochlear-implant pulse rate and inter-channel timing on channel interactions and thresholds. *J Acoust Soc Am* 116, 452-68.

- Miller, G.A., Heise, G.A., Lichten, W. 1951. The intelligibility of speech as a function of the context of the test materials. *J Exp Psychol* 41, 329-35.
- Nie, K., Barco, A., Zeng, F.G. 2006. Spectral and temporal cues in cochlear implant speech perception. *Ear Hear* 27, 208-17.
- Qin, M.K., Oxenham, A.J. 2006. Effects of introducing unprocessed low-frequency information on the reception of envelope-vocoder processed speech. *J Acoust Soc Am* 119, 2417-26.
- Rubinstein, J.T., Wilson, B.S., Finley, C.C., Abbas, P.J. 1999. Pseudospontaneous activity: stochastic independence of auditory nerve fibers with electrical stimulation. *Hear Res* 127, 108-18.
- Schatzer, R., Krenmayr, A., Au, D.K., Kals, M., Zierhofer, C. 2010. Temporal fine structure in cochlear implants: preliminary speech perception results in Cantonese-speaking implant users. *Acta Otolaryngol* 130, 1031-9.
- Seldran, F., Micheyl, C., Truy, E., Berger-Vachon, C., Thai-Van, H., Gallego, S. A model-based analysis of the “combined-stimulationadvantage.” *Hear Res*, in press.
- Seligman, P., McDermott, H.J. 1995. Architecture of the Spectra 22 speech processor. *Ann Otol Rhinol Laryngol* 104 (Suppl 166),139-41.
- Shannon, R.V. 1983. Multichannel electrical stimulation of the auditory nerve in man. II. Channel interaction. *Hear Res* 12, 1-16.
- Shannon, R.V. 1990. Forward masking in patients with cochlear implants. *J Acoust Soc Am* 88, 741-744.
- Shannon, R.V., Cruz, R.J., Galvin, J.J. 3rd. 2011. Effect of stimulation rate on cochlear implant users' phoneme, word and sentence recognition in quiet and in noise. *Audiol Neurotol*. 16, 113-23.

- Shannon, R.V., Zeng, F.G., Kamath, V., Wygonski, J., Ekelid, M. 1995. Speech recognition with primarily temporal cues. *Science* 270(5234), 303-4.
- Skinner, M.W., Arndt, P.L., Staller, S.J. 2002. Nucleus 24 advanced encoder conversion study: performance versus preference. *Ear Hear* 23(1 Suppl), 2S-17S.
- Skinner, M.W., Clark, G.M., Whitford, L.A., Seligman, P.M., Staller, S.J., Shipp, D.B., Shallop, J.K., Everingham, C., Menapace, C.M., Arndt, P.L., Antogenelli T., Brimacombe, J.A., Pijls, S., Daniels, P., George, C.R., McDermott, H.J., Beiter, A.L. 1994. Evaluation of a new spectral peak coding strategy for the Nucleus 22 Channel Cochlear Implant System. *Am J Otol* 15 (Suppl 2),15-27.
- Stickney, G.S., Loizou, P.C., Mishra, L.N., Assmann, P.F., Shannon, R.V., Opie, J.M. 2006. Effects of electrode design and configuration on channel interactions. *Hear Res* 211, 33-45.
- Studebaker, G.A. 1985. A "rationalized" arcsine transform. *J Speech Hear Res* 28, 455-62.
- Tjomov, V.L. 1971. A model to describe the results of psychoacoustical experiments on steady-state stimuli. In *Analyz Rechevikh Signalov Chelovekom*, edited by GV Gershuni (Nauka, Leningrad), 36-49.
- Tong, Y.C., Clark, G.M. 1986. Loudness summation, masking, and temporal interaction for sensations produced by electric stimulation of two sites in the human cochlea. *J Acoust Soc Am* 79, 1958-66.
- Traunmüller, H. 1990. Analytical expressions for the tonotopic sensory scale. *J Acoust Soc Am* 88, 97-100.

- Turner, C.W., Gantz, B.J., Vidal, C., Behrens, A., Henry, B.A. 2004. Speech recognition in noise for cochlear implant listeners: benefits of residual acoustic hearing. *J Acoust Soc Am* 115, 1729-35.
- Tykocinski, M., Shepherd, R.K., Clark, G.M. 1995. Reduction in excitability of the auditory nerve following electrical stimulation at high stimulus rates. *Hear Res* 88, 124-42.
- Vandali, A.E., Whitford, L.A., Plant, K.L., Clark, G.M. 2000. Speech perception as a function of electrical stimulation rate: using the Nucleus 24 cochlear implant system. *Ear Hear* 21, 608-24.
- Von Ilberg, C., Kiefer, J., Tillein, J., Pfenningdorff, T., Hartmann, R., Stürzebecher, E., Klinke, R. 1999. Electric-acoustic stimulation of the auditory system. New technology for severe hearing loss. *ORL J Otorhinolaryngol Relat Spec* 61, 334-40.
- Weber, B.P., Lai, W.K., Dillier, N., von Wallenberg, E.L., Killian, M.J., Pesch, J., Battmer, R.D., Lenarz, T. 2007. Performance and preference for ACE stimulation rates obtained with nucleus RP 8 and freedom system. *Ear Hear* 28(2 Suppl), 46S-48S.
- White, M.W., Merzenich, M.M., Gardi, J.N. 1984. Multichannel cochlear implants. Channel interactions and processor design. *Arch Otolaryngol* 110, 493-501.
- Whitford, L.A., Seligman, P.M., Everingham, C.E., Antognelli, T., Skok, M.C., Hollow, R.D., Plant, K.L., Gerin, E.S., Staller, S.J., McDermott, H.J., Gibson, W.R., Clark, G.M. 1995. Evaluation of the Nucleus Spectra 22 processor and new speech processing strategy (SPEAK) in postlinguistically deafened adults. *Acta Otolaryngol* 115, 629-37.
- Wilson, B.S. 2011. Low pulse rate cochlear implant stimulation in conjunction with a separate representation of fundamental frequencies and voiced/unvoiced distinctions. US patent US 2011/0066210 A1.

- Wilson, B.S., Finley, C.C., Farmer, J.C. Jr, Lawson, D.T., Weber, B.A., Wolford, R.D., Kenan, P.D., White, M.W., Merzenich, M.M., Schindler, R.A. 1988. Comparative studies of speech processing strategies for cochlear implants. *Laryngoscope*, 98, 1069-77.
- Wilson, B.S., Finley, C.C., Lawson, D.T., Wolford, R.D., Eddington, D.K., Rabinowitz, W.M. 1991. Better speech recognition with cochlear implants. *Nature* 352(6332), 236-8.
- Wilson, B., Wolford, R., Lawson, D. 2000a. Sixth Quarterly Progress Report. NIH project N01-DC-8-2105.
- Wilson, B., Wolford, R., Lawson, D. 2000b. Seventh Quarterly Progress Report. NIH project N01-DC-8-2105.

Figure captions

Figure 1. Effect of parameter N on percent correct in Experiment 1. Panels A-C show PC as a function of SNR, separately for the three different frame rates tested in this study. The frame rate is indicated next to each panel. For Panel D, the data were averaged across all rates. The error bars show one standard error of the mean across listeners.

Figure 2. Effect of frame rate on percent correct in Experiment 1. Panels A-D show PC as a function of SNR, separately for the four different N values tested in this study. The N value is indicated next to each panel. For Panel E, the data were averaged across all Ns. The error bars show one standard error of the mean across listeners.

Figure 3. Results of Experiment 2. Panels A and B illustrate the effect of frame rate on percent correct as a function of SNR for the different stimulation modes (A, E, and AE), for each N separately. Note that since the parameter, N, had no bearing on the A stimulation mode, the data for this stimulation mode, which are shown in these two panels, are the same; these data are plotted in both panels to facilitate comparisons of performance between the A stimulation mode and the other modes. The A and AE data are replotted in Panel C to facilitate the visual inspection of the effect of N on performance in the AE condition. The error bars show one standard error of the mean across listeners.

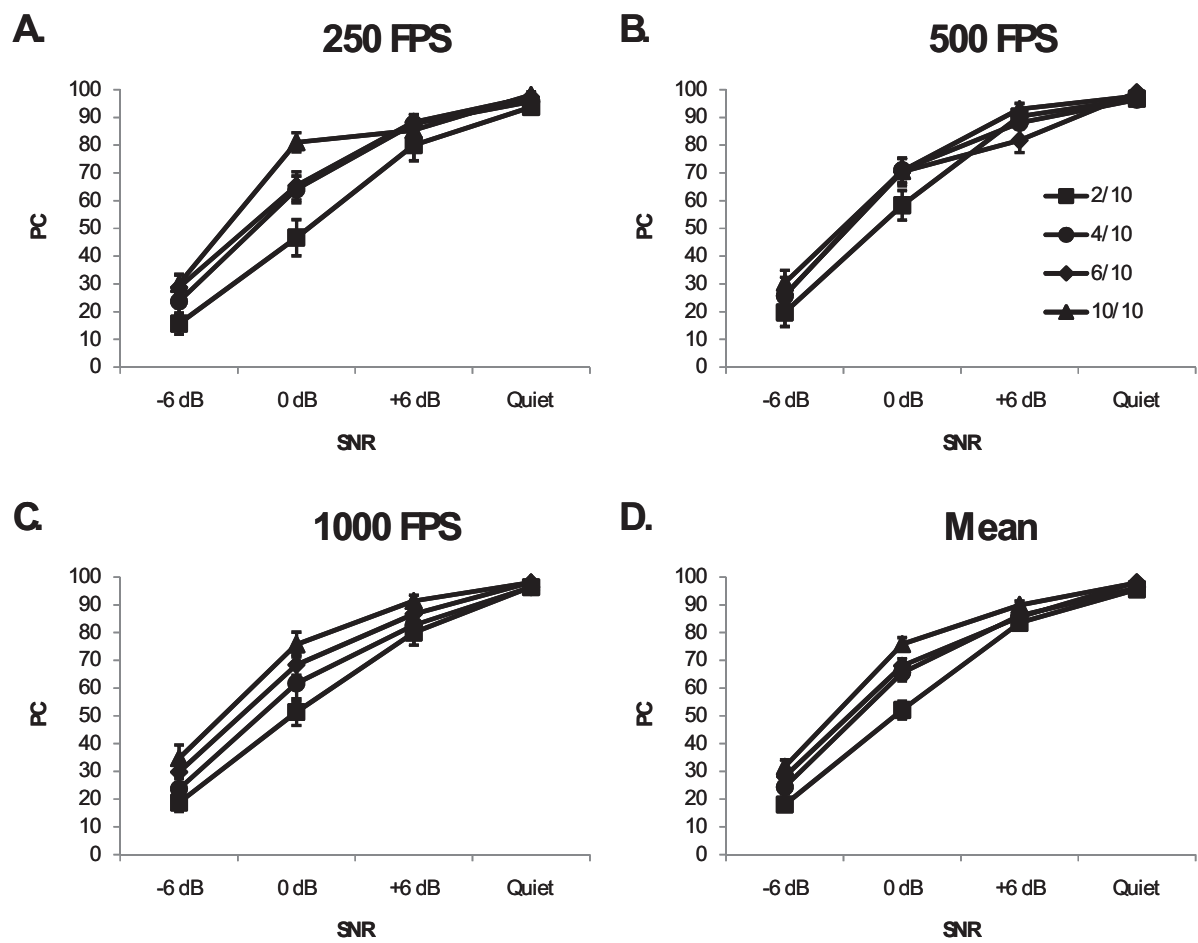


Figure 1

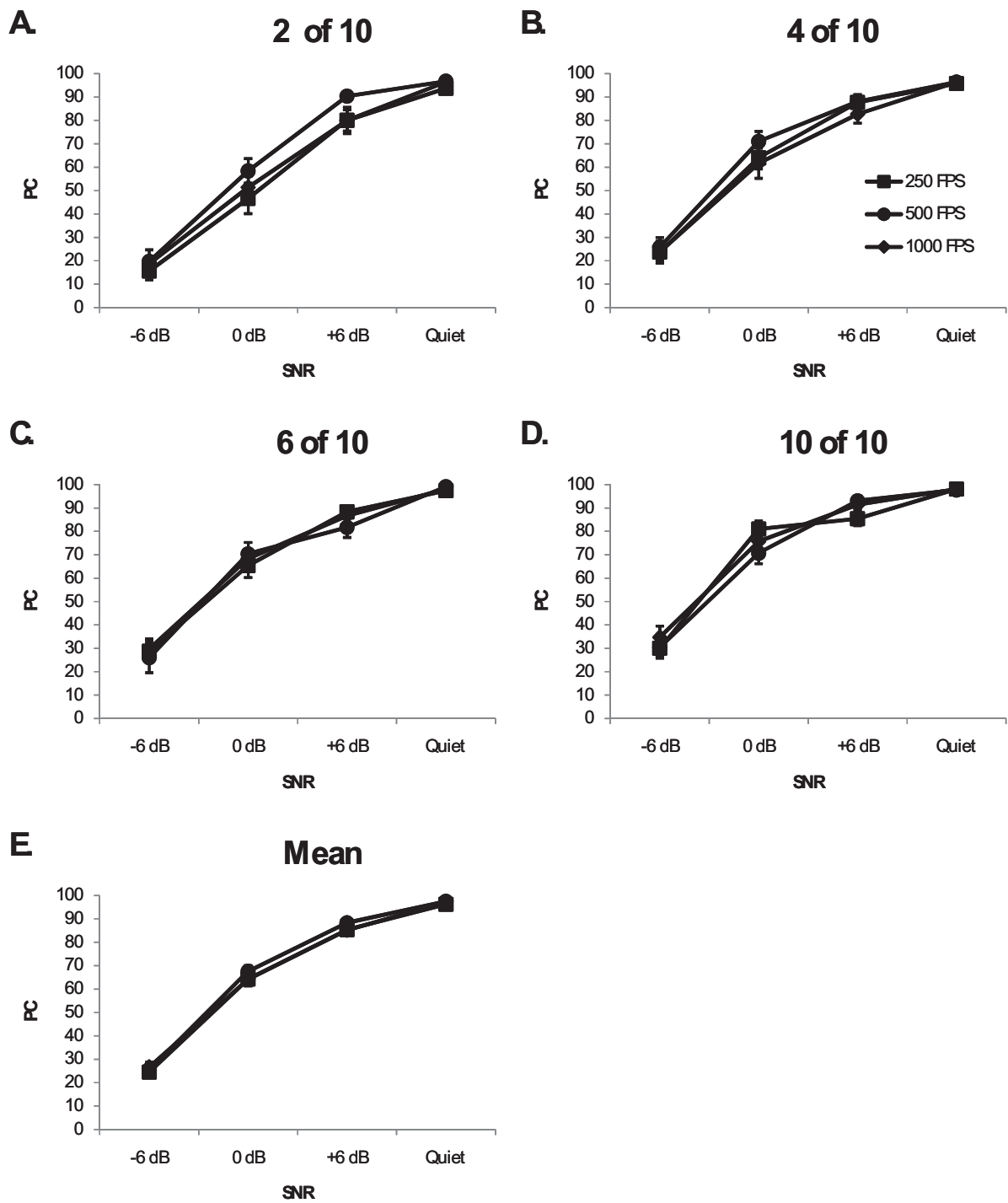


Figure 2

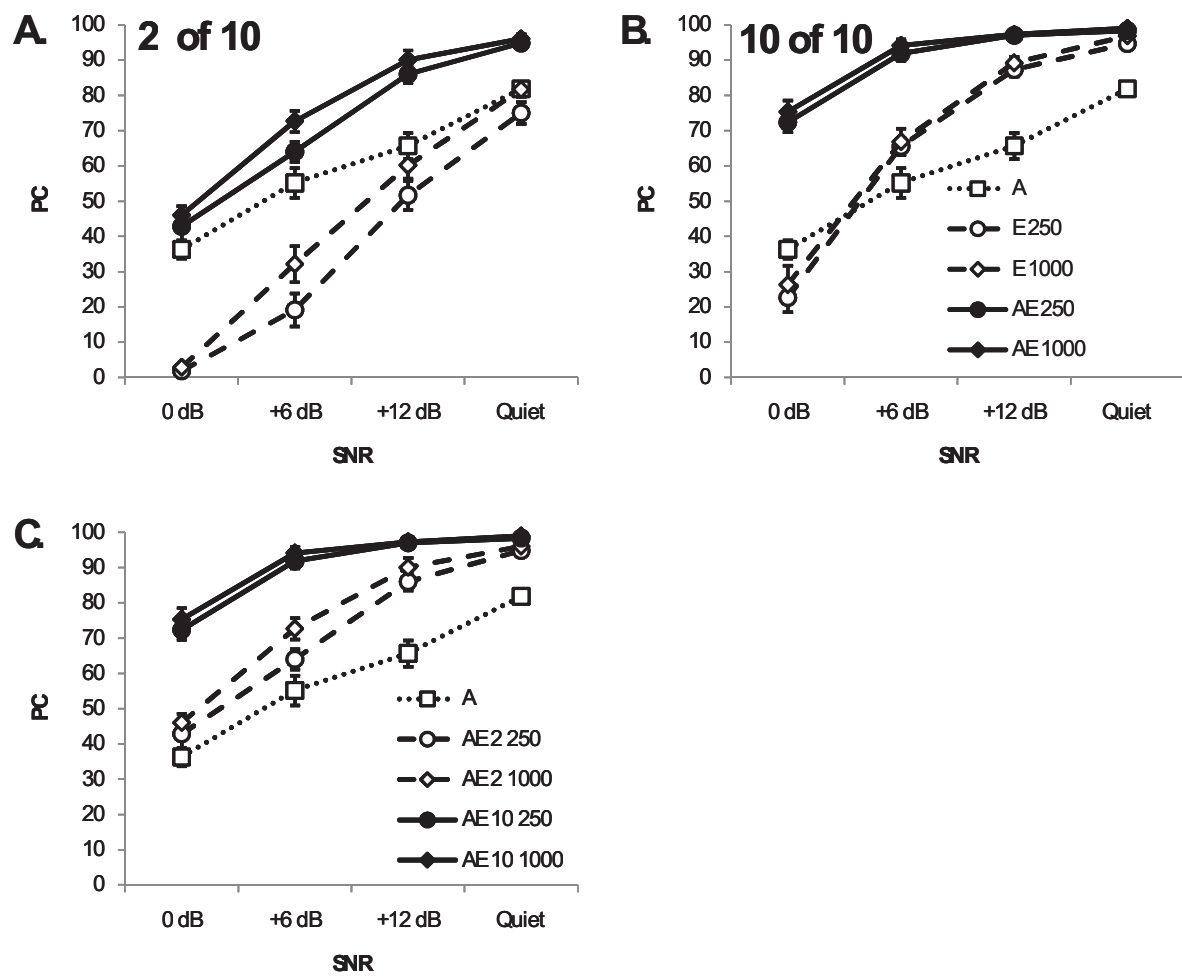


Figure 3

Table I. Center and cutoff frequencies of the vocoder analysis and synthesis bands.

Channel	Center frequency (Hz)	Center frequency (barks)	Analysis		Synthesis		Carrier
			Lower cutoff	Higher cutoff	Lower cutoff	Higher cutoff	
1	713	6.6	651	781	726	706	Sine
2	974	8.3	911	1042	986	967	Sine
3	1295	10.0	1172	1432	1247	1357	Noise
4	1687	11.7	1562	1823	1637	1748	Noise
5	2198	13.4	1953	2474	2028	2399	Noise
6	2853	15.1	2604	3125	2679	3050	Noise
7	3625	16.7	3255	4036	3330	3961	Noise
8	4600	18.3	4167	5078	4242	5003	Noise
9	5823	19.9	5208	6510	5283	6435	Noise
10	7438	21.5	6640	8333	6775	8258	Noise

III. Discussion générale

III.1. Résumé des objectifs de la thèse

Nous sommes partis du constat que, dans la littérature, tous les patients qui ont une audition basse-fréquence exploitable après l'implantation ne tirent pas un bénéfice significatif de la stimulation bimodale par rapport à leur implant seul. Pour des informations complémentaires sur les résultats détaillés des principales études sur le sujet, nous invitons le lecteur à consulter la revue de littérature de Turner et al. (2008). Dès lors nous avons proposé plusieurs hypothèses pour expliquer ces échecs et nous avons proposé de les confirmer par les différents protocoles que nous avons mis en place.

Premièrement, nous pensons que les échecs sont en partie dus au fait que les indications sont mal définies. Pour vérifier cette hypothèse, nous avons utilisé un test d'audiométrie vocale filtrée passe-bas qui nous a permis de mettre en évidence différents profils de patients. Ce test a pour but non seulement de valider si le sujet est effectivement candidat à l'implant bimodal, mais également d'aider le clinicien à choisir un mode d'appareillage plus adapté si le test venait à révéler que le sujet n'est pas un bon candidat à l'EAS.

Deuxièmement, nous pensons que les échecs peuvent être liés à une stratégie de codage pas optimisée. Jusque là, la stratégie utilisée pour l'EAS est calquée sur celle de l'implant cochléaire. Cette stratégie est efficace puisqu'elle a fait ses preuves, mais nous pensons qu'elle peut être améliorée. Ce qu'il est important de comprendre, c'est que l'implant EAS doit être considéré comme un implant à part entière : en ce sens il est important d'investiguer l'influence des différents paramètres du codage par l'implant tels que la distribution fréquentielle, le nombre de canaux, le nombre de maxima (pour les stratégies N-of-M), la cadence de stimulation ou encore la pré-emphase. De nombreux travaux ont déjà évalué l'influence de ces facteurs pour l'implant cochléaire. Nous pourrions concevoir de modifier les protocoles déjà existants pour les transposer et les appliquer à la stimulation bimodale. Certes, ces expérimentations n'auraient pas une grande originalité mais les résultats pourraient être très enrichissants pour les recommandations cliniques de l'EAS encore limitées.

III.2. Résumé des principaux résultats de la thèse

Le premier résultat qui s'est dégagé de ce manuscrit est l'existence de plusieurs profils de patients. En outre, un des résultats les plus « surprenants » est l'existence d'un profil de patients qui a des capacités supérieures à la normale pour le traitement de l'information basse fréquence. A notre connaissance, avant notre étude, le développement de capacités supra-normales chez certains patients n'était qu'une supposition (Vestergaard, 2004 : p 125) et n'avait jamais été clairement mis en évidence. Certains auteurs avaient auparavant montré de meilleures capacités chez les sourds qui présentent des ZICs (Vestergaard, 2003 ; Moore et Vinay, 2009) mais n'avaient malheureusement pas comparé les résultats avec ceux du normo-entendant. La présence et l'influence des ZICs n'a pas été investiguée dans nos travaux. D'après les critères de Moore (2001) (perte > 90 dB HL), il est clair que nos patients testés avaient des ZICs dans les hautes fréquences, mais le degré de leur surdité ne permettait pas de réaliser le TEN test car il n'aurait pas été supportable à l'intensité requise (Moore et al., 2004).

L'autre résultat surprenant qui est ressorti de nos travaux est la corrélation positive entre la durée de la surdité et les performances. Jusque là l'effet de la durée de privation n'avait été investiguée que chez l'implanté cochléaire et toutes les études de la littérature trouvaient une corrélation négative entre la durée de surdité et les performances. Il semblerait qu'en réalité, il se produise un phénomène temps-dépendant qui implique que le sourd partiel « apprend » à délaissier l'information haute fréquence à laquelle il n'a plus accès, au profit de l'information basse fréquence à laquelle il a encore accès. Après une longue durée de surdité, il y a probablement un phénomène de désafférentation des cellules hautes-fréquences au point que si un implant cochléaire est placé après cette désafférentation, le patient n'est plus capable d'intégrer normalement les informations en provenance de l'implant. Cela suppose qu'un sourd partiel qui a une surdité très ancienne ne pourra sans doute pas bénéficier de l'EAS.

Certains des sujets que nous avons testés ont été implantés en vue d'une stimulation bimodale. En accord avec ce que nous avons trouvé dans la littérature, 2 sujets sur 8 ont perdu leur audition résiduelle post-implantation. En terme de score de compréhension nous avons trouvé une grande variabilité inter-individuelle avec une évolution plus lente et des scores moins bons chez les sujets qui avaient une surdité ancienne. Notre 1^{er} sujet implanté EAS a par exemple gardé ses restes auditifs post-implantation, mais il obtient de mauvais résultats de

compréhension, probablement parce qu'il entend encore jusqu'à 2 kHz. Aujourd'hui, avec les connaissances que nous avons acquises au cours de nos travaux, nous n'implanterions pas un tel candidat.

Une autre partie de nos travaux a consisté à réaliser des simulations d'implant EAS. Nous avons ainsi montré que les performances obtenues en condition EAS pouvaient être prédites à partir des scores obtenus en conditions de stimulation acoustique seule et électrique seule, grâce à des modèles de combinaison de l'information. Nous avons en particulier montré que tous les modèles ne sont pas aussi précis pour prédire les résultats en EAS et que le phénomène de « superadditivité » observé par Kong et Carlyon (2007) pouvait être simplement expliqué par un modèle plus optimal.

Enfin, d'autres travaux par simulation suggèrent que des stratégies de codage qui limitent les interactions en réduisant la quantité d'information instantanée ne nuisent pas nécessairement à la compréhension en EAS. Ces observations doivent désormais être validées chez l'implanté EAS. Ce sujet sera développé dans la section suivante.

III.3. Perspectives

* Optimisation du bilan pour améliorer le choix de l'appareillage

Nous avons mis en évidence, par l'utilisation de tests audiométriques avancés, qu'il existait différents profils de patients. D'après nos observations, nous pensons que tous les patients ne sont pas candidats à l'EAS, bien qu'ils entrent initialement dans les critères d'implantation d'après leur audiométrie tonale et vocale. D'ailleurs notre 1^{er} patient implanté EAS avait en réalité un mauvais profil ; il obtient de mauvais résultats alors que son audiométrie tonale est bien préservée.

Pour restituer les hautes fréquences dans le cas de ces surdités partielles, il existe plusieurs alternatives d'appareillages, à savoir les appareils conventionnels par amplification, les appareils à compression fréquentielle, les implants d'oreille moyenne, la stimulation électro-acoustique et l'implant cochléaire.

Chacun de ces modes d'appareillage a des indications bien particulières, mais dans le cas des surdités partielles, les indications des différents systèmes se chevauchent, elles sont à ce jour mal définies et le sujet est encore peu exploré. Si nous regardons les résultats obtenus avec ces systèmes, tous présentent des succès et des échecs et les données individuelles sont très variables. Aujourd'hui, la vraie question est de comprendre pourquoi ces systèmes fonctionnent chez certains sujets et pas chez d'autres. L'outil d'audiométrie vocale filtrée que nous avons utilisé a mis en évidence qu'il existait plusieurs profils de patients. En particulier, nous avons montré que l'âge d'apparition de la surdité et la durée de privation avaient une influence majeure et affectaient les performances, beaucoup plus que l'audibilité. Nous pensons qu'il est important, dans les prochaines années, que des outils complémentaires soient mis en place afin d'optimiser le choix de l'appareillage et de limiter les échecs.

* Amélioration de la technique chirurgicale

Pour que les patients puissent bénéficier de la stimulation bimodale, il est indispensable de conserver l'audition résiduelle après la chirurgie. Depuis une quinzaine d'années, nous observons un développement des chirurgies atraumatiques et mini-invasives. Deux facteurs ont contribué à ce développement. Le premier facteur est l'amélioration des techniques d'implantation. Nous trouvons par exemple des recommandations pour les insertions par la fenêtre ronde versus par cochléostomie, et des recommandations pour l'utilisations de drogues et pour l'utilisations de cathéters qui permettent de délivrer ces drogues (Paasche et al., 2003 ; Richardson et al., 2006 ; Ibrahim et al., 2011). Le second facteur est une meilleure conception des porte-électrodes (par exemple les électrodes Flex de Med-El ou le stylet et la technique AOS chez Cochlear), qui permettent de préserver et de ne pas endommager les structures fragiles de la cochlée.

De nombreuses études ont montré qu'il était possible de conserver l'audition résiduelle dans la majorité des cas (Gstoettner et al., 2006 ; James et al., 2006) avec des insertions peu profondes de l'ordre de 360-400° maximum. Des études récentes ont mis en évidence que la préservation de l'audition était un objectif réalisable avec un taux de réussite supérieur à 80%, que ce soit pour des portes-électrodes courts (Helbig et al., 2011a avec l'électrode Flex EAS), mais aussi et surtout avec des porte-électrodes longs et des insertions profondes qui couvrent parfois la zone où l'audition est encore partiellement saine (Prentiss et al., 2010 ; Helbig et al., 2011b ; Usami et al., 2011), ce qui n'était pas forcément concevable il y a 15 ou 20 ans.

L'objectif dans les prochaines années sera de mieux maîtriser les mécanismes responsables de la perte d'audition post-implantation et de limiter le risque de perdre l'audition résiduelle après la chirurgie.

*** Extension des indications de l'EAS et optimisation du choix du porte-électrode**

Grâce aux techniques qui permettent de préserver l'audition aussi bien pour les insertions courtes que profondes, il sera possible, probablement dans un futur proche, de mieux sélectionner le type d'électrode en fonction de l'audition résiduelle. Par exemple, avec les électrodes Med-El, pour un patient qui a des reliquats d'audition jusqu'à 1000 Hz, nous nous orienterons vers une électrode courte type Flex EAS (24 mm) ou Flex 20 (en développement chez Med-El), s'il a des reliquats d'audition jusqu'à 250 Hz ou moins, nous lui proposerons une électrode longue type Flex Soft (31 mm) et s'il a des reliquats d'audition jusqu'à 500-700 Hz, nous proposerons une électrode longue type Flex Soft insérée partiellement ou une électrode de longueur intermédiaire type Flex 28 (28 mm) qui est actuellement en développement chez Med-El.

<http://www.medel.com/data/downloads/MAESTRO/23236.pdf>

*** Intérêt d'étendre la plage de stimulation acoustique**

Les processeurs disponibles actuellement pour l'EAS ont des écouteurs similaires aux aides auditives conventionnelles, qui ne véhiculent pas parfaitement les sons de très basse fréquence. La courbe de réponse des haut-parleurs est souvent atténuée pour les sons de fréquence inférieure à 150-200 Hz. Or il a été montré, en particulier en condition de bruit, que l'information acoustique jusqu'à 125 Hz lorsqu'elle est combinée à la stimulation de l'implant améliore significativement les performances par rapport à la condition implant seul (Zhang et al., 2010). Cette information acoustique présentée seule ne fournit aucune intelligibilité, mais couplée à l'implant, elle permet de transmettre des indices acoustiques tels que le F_0 et les transitions des premiers formants qui aident à la ségrégation de la parole dans le bruit (Başkent & Chatterjee, 2010; Brown & Bacon, 2009a, b, 2010; Kong et al., 2005; Qin & Oxenham, 2006; Turner et al., 2004) et qui sont particulièrement robustes en milieu bruyant (Alwan et al., 2011). On peut supposer que les firmes qui proposent l'EAS ont tout intérêt à développer des systèmes avec des écouteurs capables de transmettre cette information très basse fréquence (< 250 Hz voire même < 100 Hz). Pour encourager les fabricants, il faudrait

mettre en évidence par des expérimentations complémentaires à celle de Zhang et al. (2010), et avec des populations plus importantes, que l'information très basse fréquence fournie acoustiquement peut améliorer les performances de l'EAS dans les conditions d'écoute difficile.

* Décalage tonotopique : expérience avec l'IC / quelles conséquences pour l'EAS ?

La réhabilitation de la surdité par l'implant cochléaire est assez complexe et nécessite une rééducation qui peut parfois durer plusieurs mois voire plusieurs années avant d'obtenir une compréhension optimale (Tyler et al., 1997 ; Wilson et Dorman, 2008). Ceci s'explique notamment par le fait que l'implant nécessite une réattribution fréquentielle et on peut supposer que la profondeur d'insertion de l'électrode va avoir une influence sur l'évolution plus ou moins rapide des performances et sur la durée de la rééducation.

Concernant l'effet de la profondeur d'insertion sur l'intelligibilité avec l'implant, toutes les études qui ont investigué ce sujet chez l'implanté cochléaire ne présentent pas des résultats formels. Par exemple en 2002, l'équipe de Skinner et al. avait montré que les performances étaient positivement corrélées à la profondeur d'insertion du porte-électrode ($r^2 = 0.17$, $p = 0.04$ lorsque la profondeur d'insertion est exprimée en mm ; $r^2 = 0.20$, $p = 0.02$ lorsque la profondeur est exprimée en pourcentage de la longueur cochléaire). Mais une étude plus récente de Lee et al. (2010) a trouvé au contraire que les performances n'étaient pas nécessairement liées à la profondeur d'insertion ($r^2 = 0.05$, $p = 0.27$), ceci peut probablement s'expliquer par des stratégies de codage plus évoluées car plus récentes dans l'étude de Lee et al. et qui pourraient minimiser l'effet de la profondeur d'insertion sur les performances.

En parallèle, des équipes ont réalisé chez l'implanté cochléaire des expérimentations avec des tâches de comparaison de hauteur tonale acoustique versus électrique. Généralement les sujets avaient une audition résiduelle ou une audition normale du côté non implanté et avaient pour consigne de comparer la hauteur tonale de l'électrode la plus apicale (ou de plusieurs électrodes) avec une stimulation acoustique, généralement un son pur, présentée à l'oreille controlatérale à l'implant. L'ensemble des études montre, comme on peut s'y attendre, que la sensation de hauteur est plus grave pour les insertions plus profondes (Blamey et al., 1996 ; Boëx et al., 2006 ; Dorman et al., 2007 ; Vermeire et al., 2008 ; Carlyon et al., 2010 ; Baumann et al., 2011) et cette sensation de hauteur serait plus grave d'une à deux octave par

rapport à ce qui peut être attendu sur la base de l'équation de Greenwood (1990) (Boëx et al., 2006 ; Dorman et al., 2007 ; McDermott et al., 2009). En réalité chez l'implanté, il se produit une acclimatation voire une plasticité qui va engendrer une variation de la sensation de hauteur tonale avec le temps (figure 24) au point qu'un implanté peut voir sa hauteur se décaler de une à deux octave vers les fréquences graves au cours des deux premières années d'utilisation de l'implant (Reiss et al., 2007 ; McDermott et al., 2009).

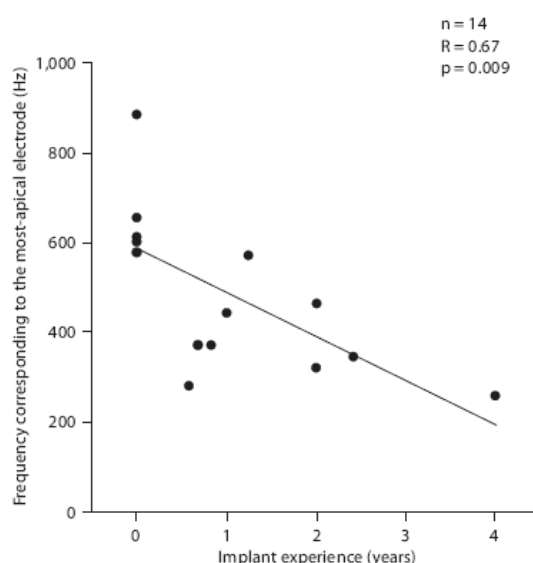


Figure 24. Fréquence acoustique correspondant à la hauteur produite par une stimulation de l'électrode la plus apicale, en fonction du temps d'utilisation de l'implant. D'après McDermott et al., 2009.

Certains travaux par simulation ont suggéré que les implantés pouvaient s'adapter avec le temps et apprendre à comprendre avec une carte décalée vers les aigus (Rosen et al., 1999 ; Fu et al., 2005), mais les travaux de Faulkner (2006) et de Baskent et Shannon (2007) ont montré qu'un décalage tonotopique, qu'il soit apical ou basal, nuisait tout de même à la compréhension et pouvait sévèrement affecter les performances des patients implantés cochléaires et EAS.

D'après Reiss et al. (2007), dans le cas d'une insertion courte (implant hybrid, 10 mm, 6 électrodes) pour l'EAS, les performances avec l'implant au-delà de un an d'utilisation seraient positivement corrélées à la sensation de hauteur mesurée avant 3 mois d'utilisation de

l'implant ($r^2 = 0.46$; $p < 0.05$). Ce qui signifie, d'après l'interprétation des auteurs, que si la sensation de hauteur est haute les premiers jours de port de l'implant, alors le capital neuronal de la base de la cochlée est probablement riche, et la sensation de hauteur perçue correspondra à peu près à la tonotopie naturelle. Au contraire, si la hauteur tonale est perçue plus grave, c'est sans doute que le capital neuronal basal est pauvre et que ce sont des cellules plus apicales qui vont répondre à la stimulation (stimulation « ectopique »). En général, il semble donc important que la tonotopie naturelle soit respectée.

A titre d'exemple une étude de Hochmair et al. (2003) montre que 8 canaux de stimulation répartis sur l'intégralité de la cochlée fournissent de meilleurs résultats dans le silence et dans le bruit que 8 canaux qui seraient répartis sur le tour basal et milieu de la cochlée. L'utilisation d'électrodes longues permet ainsi de minimiser ce phénomène de décalage puisque les lieux de stimulation de l'implant respectent plus les lieux d'excitation naturels. D'autres travaux conduits chez des implantés avec des restes auditifs bilatéraux, ont montré une corrélation négative entre le décalage tonotopique et les performances (respectivement $r = -0.87$ et $r = -0.64$, $p < 0.05$ pour l'identification des voyelles et des consonnes : Di Nardo et al., 2008) et dans certains cas, une réattribution fréquentielle ou une désactivation des électrodes les plus apicales (où il y a de l'audition résiduelle) peut s'avérer utile pour améliorer les performances (Gani et al., 2007 ; Di Nardo et al., 2008).

Qu'en est-il pour l'EAS ?

La profondeur d'insertion est justement une des interrogations majeures pour l'électro-acoustique et le sujet est très discuté. Faut-il une insertion courte ? une insertion longue ? une insertion fonction de l'audition résiduelle ? Et les questions qui en découlent : faut-il un chevauchement ou un léger trou fréquentiel entre les stimulations acoustique et électrique ? Pour le codage par l'implant, faut-il utiliser la même distribution fréquentielle qu'un implant cochléaire ? ou bien faut-il respecter la distribution naturelle en fonction de la profondeur d'insertion de l'électrode afin de mieux fusionner avec l'information acoustique ?

Une insertion courte présente l'avantage de limiter le traumatisme aux cellules basales. Cependant il existe une contrepartie non négligeable à cette insertion courte, à savoir que le

patient risque d'être beaucoup plus handicapé (par rapport à un implanté cochléaire), dans la mesure où il risque d'y avoir une distorsion importante en terme de place tonotopique (Gantz et Turner, 2003 ; Woodson et al., 2010). A cause de ces distorsions, il faut compter plusieurs mois avant que les patients EAS apprennent à comprendre avec la stimulation bimodale (Gantz et Turner, 2003).

Pour remédier à cela, nous pensons qu'il faudrait utiliser des porte-électrodes adaptés à l'audiogramme de chaque patient. Aujourd'hui, il est possible, avec des insertions profondes (électrode Flex soft Medel) de préserver l'audition dans les mêmes proportions de réussite qu'une insertion courte (Helbig et al., 2011b ; Usami et al., 2011). Selon Nadol et al. (1994), le dommage des structures cochléaires se produit en particulier sur les 15 premiers mm de la base et une insertion plus profonde n'augmenterait pas nécessairement le risque de trauma cochléaire (Gstoettner et al., 1997), ce qui expliquerait les résultats comparables pour les insertions courtes et longues. L'utilisation d'un porte-électrode fonction de l'audiogramme du patient n'est donc pas inconcevable (cf section précédente sur l'optimisation du choix du porte-électrode) et devrait permettre de mieux maîtriser le décalage fréquentiel.

Un axe de recherche à explorer ces prochaines années sera de comprendre et de minimiser le décalage tonotopique. Selon Stakhovskaya et al. (2007), la hauteur associée à la position de l'électrode, immédiatement après l'implantation est souvent proche de ce qui est prédit par la carte fréquentielle du ganglion spiral. Aussi, d'après Reiss et al. (2007), ce sont les sensations de hauteur très précoces qui vont être les meilleurs prédicteurs des performances à long terme. On peut donc supposer que si dès l'activation de l'implant, la distribution fréquentielle respecte la tonotopie naturelle d'après la profondeur d'insertion du porte-électrode, alors le travail de rééducation sera certainement moins difficile et les performances à long terme seront probablement meilleures que si la distribution est décalée vers les graves ou vers les aigus.

* Optimisation de l'implant et de la stratégie de codage

Nos travaux par simulation ont suggéré que lorsque les patients avaient des reliquats d'audition dans les basses fréquences, il était possible de réduire le phénomène d'interaction 1/ soit en réduisant le nombre de canaux, 2/ soit en utilisant une stratégie à extraction de pics, 3/ soit en réduisant la cadence de stimulation.

1/ Limitation du nombre de canaux

Si dans les prochaines années les longueurs des porte-électrodes peuvent être adaptées à l'audition résiduelle, alors pour les patients qui ont des restes auditifs jusqu'à 1000-1400 Hz, il suffira probablement d'un implant monocanal ou bien à deux ou trois canaux de stimulation pour restituer les hautes fréquences. Sur un porte-électrode court de l'ordre de 20 mm, l'espace inter-électrode sera augmenté, ce qui diminuera le phénomène d'interactions, par rapport à un porte-électrode de même longueur qui dispose actuellement de 12 canaux de stimulation et qui a des contacts plus proches. Encore une fois, ceci ne sera possible que lorsque la chirurgie sera mieux maîtrisée.

2/ Utilisation d'une stratégie à extraction de pics

Supposons encore une fois que dans l'avenir, la longueur de l'électrode soit fonction de l'audition résiduelle. Plaçons nous cette fois dans la condition où le patient n'a des reliquats d'auditions que jusqu'à 250-500 Hz et dans ce cas nous utiliserons une électrode longue. Dans ce cas de figure, nous ne chercherons pas à réduire nécessairement le nombre de canaux sur le porte-électrode, mais plutôt à diminuer le nombre de canaux stimulés à chaque instant (trame) de stimulation. Nous pensons que dans une condition de $N\text{-of-}M = 2/10$ canaux stimulés avec une pré-emphase appropriée, les performances dans le bruit peuvent être sensiblement proches d'une condition 10/10 où tous les canaux sont stimulés. Ce qui signifie que les performances ne sont que légèrement affectées alors que le nombre de stimulations (i.e. d'informations) par seconde est pourtant divisé par 5 (d'où une réduction des interactions).

3/ Réduction de la cadence de stimulation

Dans l'implant cochléaire, il a été montré que des cadences de stimulation rapides fournissaient une meilleure représentation de l'information temporelle (Wilson et al., 1991; Rubinstein et al., 1999). Mais les cadences de stimulation rapides sont généralement responsables d'interactions inter-canaux (De Balthasar et al., 2003, Shannon, 1990, Tykocinski et al., 1995; Middlebrooks, 2004) et dans le cas de l'EAS, nous pouvons nous demander si l'utilisation de cadences si rapides est bien utile.

Pour l'EAS, un brevet récent (Wilson, 2011) suggère d'utiliser une cadence de stimulation très lente pour les composantes haute fréquence véhiculées par l'implant. Selon Wilson il est possible de stimuler à des cadences exceptionnellement basses dans la mesure où la fréquence fondamentale, le premier ou les deux premiers harmoniques ainsi que les distinctions de périodicité sont déjà très bien représentés par la stimulation acoustique. Toujours d'après Wilson, les cadences de stimulation très lentes permettent de réduire les interactions inter-électrodes et d'améliorer la sensibilité à la modulation, par rapport à des cadences plus rapides, ce qui permet une meilleure perception de la parole. C'est pourquoi nous pensons qu'il est déterminant d'évaluer l'influence de la cadence de stimulation sur les performances de l'implanté EAS.

En définitif, en dehors des études de cas sur la technique chirurgicale, la préservation de l'audition et les performances avec l'EAS, la recherche sur la stimulation bimodale est aujourd'hui assez pauvre.

Ce qu'il est important de comprendre, c'est que l'implant bimodal est un implant spécifique, à part entière et qu'au niveau des recommandations cliniques, ce qui est vrai pour l'implant cochléaire n'est pas nécessairement vrai pour l'implant EAS.

IV. Références bibliographiques

- Adunka O, Kiefer J, Unkelbach MH, Lehnert T, Gstoettner W. Development and evaluation of an improved cochlear implant electrode design for electric acoustic stimulation. *Laryngoscope*. 2004;114(7):1237-41.
- Alwan A, Jiang J, Chen W. Perception of Place of Articulation for Plosives and Fricatives in Noise. *Speech Commun*. 2011;53(2):195-209.
- Amos NE, Humes LE. Contribution of high frequencies to speech recognition in quiet and noise in listeners with varying degrees of high-frequency sensorineural hearing loss. *J Speech Lang Hear Res*. 2007;50(4):819-34.
- Arnoldner C, Helbig S, Wagenblast J, Baumgartner WD, Hamzavi JS, Riss D, Gstoettner W. Electric acoustic stimulation in patients with postlingual severe high-frequency hearing loss: clinical experience. *Adv Otorhinolaryngol*. 2010;67:116-24.
- Baer T, Moore BC. Effects of spectral smearing on the intelligibility of sentences in the presence of noise. *J Acoust Soc Am*. 1993; 94(3), 1229–1241.
- Baer T, Moore BC, Kluk K. Effects of low pass filtering on the intelligibility of speech in noise for people with and without dead regions at high frequencies. *J Acoust Soc Am*. 2002;112(3 Pt 1):1133-44.
- Başkent D, Chatterjee M. Recognition of temporally interrupted and spectrally degraded sentences with additional unprocessed low-frequency speech. *Hear Res*. 2010 1;270(1-2):127-33.
- Başkent D, Shannon RV. Combined effects of frequency compression-expansion and shift on speech recognition. *Ear Hear*. 2007;28(3):277-89.
- Baumann U, Rader T, Helbig S, Bahmer A. Pitch matching psychometrics in electric acoustic stimulation. *Ear Hear*. 2011;32(5):656-62.

- Blamey PJ, Dooley GJ, Parisi ES, Clark GM. Pitch comparisons of acoustically and electrically evoked auditory sensations. *Hear Res.* 1996;99(1-2):139-50.
- Boeheim K, Pok SM, Schloegel M, Filzmoser P. Active middle ear implant compared with open-fit hearing aid in sloping high-frequency sensorineural hearing loss. *Otol Neurotol.* 2010;31(3):424-9.
- Boëx C, Baud L, Cosendai G, Sigrist A, Kós MI, Pelizzone M. Acoustic to electric pitch comparisons in cochlear implant subjects with residual hearing. *J Assoc Res Otolaryngol.* 2006;7(2):110-24.
- Briggs RJ, Tykocinski M, Stidham K, Roberson JB. Cochleostomy site: implications for electrode placement and hearing preservation. *Acta Otolaryngol.* 2005;125(8):870-6.
- Brown CA, Bacon SP. Low-frequency speech cues and simulated electric-acoustic hearing. *J Acoust Soc Am.* 2009a;125(3):1658-1665.
- Brown CA, Bacon SP. Achieving electric-acoustic benefit with a modulated tone. *Ear Hear.* 2009b;30(5):489-493.
- Brown CA, Bacon SP. Fundamental frequency and speech intelligibility in background noise. *Hear Res.* 2010;266(1-2):52-59.
- Carlyon RP, Macherey O, Frijns JH, Axon PR, Kalkman RK, Boyle P, Baguley DM, Briggs J, Deeks JM, Briare JJ, Barreau X, Dauman R. Pitch comparisons between electrical stimulation of a cochlear implant and acoustic stimuli presented to a normal-hearing contralateral ear. *J Assoc Res Otolaryngol.* 2010;11(4):625-40.
- Chang JE, Bai JY, Zeng FG. Unintelligible low-frequency sound enhances simulated cochlear-implant speech recognition in noise. *IEEE Trans Biomed Eng.* 2006;53(12 Pt 2):2598-601.

- Ching TY, Dillon H, Byrne D. Speech recognition of hearing-impaired listeners: predictions from audibility and the limited role of high-frequency amplification. *J Acoust Soc Am*. 1998;103(2):1128-40.
- Cohen LT, Saunders E, Knight MR, Cowan RS. Psychophysical measures in patients fitted with Contour and straight Nucleus electrode arrays. *Hear Res*. 2006;212(1-2):160-75.
- de Balthasar C, Boëx C, Cosendai G, Valentini G, Sigrist A, Pelizzone M. Channel interactions with high-rate biphasic electrical stimulation in cochlear implant subjects. *Hear Res*. 2003;182:77-87.
- Di Nardo W, Cantore I, Marchese MR, Cianfrone F, Scorpecci A, Giannantonio S, Paludetti G. Electric to acoustic pitch matching: a possible way to improve individual cochlear implant fitting. *Eur Arch Otorhinolaryngol*. 2008;265:1321–1328.
- Dorman MF, Gifford RH, Spahr AJ, McKarns SA. The benefits of combining acoustic and electric stimulation for the recognition of speech, voice and melodies. *Audiol Neurotol*. 2008;13(2):105-12.
- Dorman MF, Spahr T, Gifford R, Loiselle L, McKarns S, Holden T, Skinner M, Finley C. An electric frequency-to-place map for a cochlear implant patient with hearing in the nonimplanted ear. *J Assoc Res Otolaryngol*. 2007;8(2):234-40.
- Driscoll CL, Carlson ML, Fama AF, Lane JJ. Evaluation of the hybrid-L24 electrode using microcomputed tomography. *Laryngoscope*. 2011;121(7):1508-16.
- Dunn CC, Tyler RS, Witt SA. Benefit of wearing a hearing aid on the unimplanted ear in adult users of a cochlear implant. *J Speech Lang Hear Res*. 2005;48(3):668-80.
- Eshraghi AA, Frachet B, Van De Water TR, Eter E. Surdit  de l'adulte : r parer ou compenser. *Rev Prat*. 2009;59:645-52.

- Faulkner A. Adaptation to distorted frequency-to-place maps: implications of simulations in normal listeners for cochlear implants and electroacoustic stimulation. *Audiol Neurotol.* 2006;11 Suppl 1:21-6.
- Fitzgerald MB, Sagi E, Jackson M, Shapiro WH, Roland JT Jr, Waltzman SB, Svirsky MA. Reimplantation of hybrid cochlear implant users with a full-length electrode after loss of residual hearing. *Otol Neurotol.* 2008;29(2):168-73.
- Florentine M, Buus S, Scharf B, Zwicker E. Frequency selectivity in normally-hearing and hearing-impaired observers. *J Speech Hear Res.* 1980;23(3):646-69.
- Frayse B, Lavieille JP, Schmerber S, Enée V, Truy E, Vincent C, Vanecloo FM, Sterkers O. A multicenter study of the Vibrant Soundbridge middle ear implant: early clinical results and experience. *Otol Neurotol.* 2001;22(6):952-61.
- Frayse B, Macías AR, Sterkers O, Burdo S, Ramsden R, Deguine O, Klenzner T, Lenarz T, Rodriguez MM, Von Wallenberg E, James C. Residual hearing conservation and electroacoustic stimulation with the nucleus 24 contour advance cochlear implant. *Otol Neurotol.* 2006;27(5):624-33.
- Fu QJ, Shannon RV. Effects of electrode location and spacing on phoneme recognition with the Nucleus-22 cochlear implant. *Ear Hear.* 1999;20(4):321-31.
- Füllgrabe C, Meyer B, Lorenzi C. Effect of cochlear damage on the detection of complex temporal envelopes. *Hear Res.* 2003;178(1-2):35-43.
- Gani M, Valentini G, Sigrist A, Kós MI, Boëx C. Implications of deep electrode insertion on cochlear implant fitting. *J Assoc Res Otolaryngol.* 2007;8(1):69-83.
- Gantz BJ, Hansen MR, Turner CW, Oleson JJ, Reiss LA, Parkinson AJ. Hybrid 10 clinical trial: preliminary results. *Audiol Neurotol.* 2009;14 Suppl 1:32-8.
- Gantz BJ, Turner CW. Combining acoustic and electrical hearing. *Laryngoscope.* 2003;113(10):1726-30.

- Gantz BJ, Turner C, Gfeller KE. Acoustic plus electric speech processing: preliminary results of a multicenter clinical trial of the Iowa/Nucleus Hybrid implant. *Audiol Neurotol.* 2006;11 Suppl 1:63-8.
- Gantz BJ, Turner C, Gfeller KE, Lowder MW. Preservation of hearing in cochlear implant surgery: advantages of combined electrical and acoustical speech processing. *Laryngoscope.* 2005;115(5):796-802.
- Glasberg BR, Moore BC. Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *J Acoust Soc Am.* 1986;79(4):1020-33.
- Glasberg BR, Moore BC. Derivation of auditory filter shapes from notched-noise data. *Hear Res.* 1990 1;47(1-2):103-38.
- Greenwood DD. A cochlear frequency-position function for several species--29 years later. *J Acoust Soc Am.* 1990;87(6):2592-605.
- Gstoettner WK, Baumgartner WD, Franz P, Hamzavi J. Cochlear implant deep-insertion surgery. *Laryngoscope.* 1997;107(4):544-6.
- Gstoettner WK, Helbig S, Maier N, Kiefer J, Radeloff A, Adunka OF. Ipsilateral electric acoustic stimulation of the auditory system: results of long-term hearing preservation. *Audiol Neurotol.* 2006;11 Suppl 1:49-56.
- Gstoettner W, Helbig S, Settevendemie C, Baumann U, Wagenblast J, Arnoldner C. A new electrode for residual hearing preservation in cochlear implantation: first clinical results. *Acta Otolaryngol.* 2009;129(4):372-9.
- Gstoettner W, Kiefer J, Baumgartner WD, Pok S, Peters S, Adunka O. Hearing preservation in cochlear implantation for electric acoustic stimulation. *Acta Otolaryngol.* 2004;124(4):348-52.

- Gstoettner WK, van de Heyning P, O'Connor AF, Morera C, Sainz M, Vermeire K, McDonald S, Cavallé L, Helbig S, Valdecasas JG, Anderson I, Adunka OF. Electric acoustic stimulation of the auditory system: results of a multi-centre investigation. *Acta Otolaryngol.* 2008;128(9):968-75.
- Hamzavi J, Pok SM, Gstoettner W, Baumgartner WD. Speech perception with a cochlear implant used in conjunction with a hearing aid in the opposite ear. *Int J Audiol.* 2004;43(2):61-5.
- Heinz MG, Swaminathan J, Boley JD, Kale S. Across-Fiber Coding of Temporal Fine-Structure: Effects of Noise-Induced Hearing Loss on Auditory-Nerve Responses. *The Neurophysiological Bases of Auditory Perception 2010, Part 9*, 621-630
- Helbig S, Baumann U, Hey C, Helbig M. Hearing preservation after complete cochlear coverage in cochlear implantation with the free-fitting FLEXSOFT electrode carrier. *Otol Neurotol.* 2011;32(6):973-9.
- Helbig S, Van de Heyning P, Kiefer J, Baumann U, Kleine-Punte A, Brockmeier H, Anderson I, Gstoettner W. Combined electric acoustic stimulation with the PULSARCI(100) implant system using the FLEX(EAS) electrode array. *Acta Otolaryngol.* 2011a;131(6):585-95.
- Hochmair I, Arnold W, Nopp P, Jolly C, Müller J, Roland P. Deep electrode insertion in cochlear implants: apical morphology, electrodes and speech perception results. *Acta Otolaryngol.* 2003;123(5):612-7.
- Hodges AV, Schloffman J, Balkany T. Conservation of residual hearing with cochlear implantation. *Am J Otol.* 1997;18(2):179-83.
- Hogan CA, Turner CW. High-frequency audibility: benefits for hearing-impaired listeners. *J Acoust Soc Am.* 1998;104(1):432-41.

- Hopkins K, Moore BC, Stone MA. Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech. *J Acoust Soc Am*. 2008;123(2):1140-53.
- Hopkins K, Moore BC. Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information. *J Acoust Soc Am*. 2007;122(2):1055-68.
- Hornsby BW, Ricketts TA. The effects of hearing loss on the contribution of high- and low-frequency speech information to speech understanding. *J Acoust Soc Am*. 2003;113(3):1706-17.
- Hornsby BW, Ricketts TA. The effects of hearing loss on the contribution of high- and low-frequency speech information to speech understanding. II. Sloping hearing loss. *J Acoust Soc Am*. 2006;119(3):1752-63.
- Horwitz AR, Ahlstrom JB, Dubno JR. Factors affecting the benefits of high-frequency amplification. *J Speech Lang Hear Res*. 2008;51(3):798-813.
- Huang CQ, Shepherd RK. Reduction in excitability of the auditory nerve following electrical stimulation at high stimulus rates: V. Effects of electrode surface area. *Hear Res*. 2000;146(1-2):57-71.
- Hughes ML, Abbas PJ. Electrophysiologic channel interaction, electrode pitch ranking, and behavioral threshold in straight versus perimodiolar cochlear implant electrode arrays. *J Acoust Soc Am*. 2006;119(3):1538-47.
- Ibrahim HN, Bossard D, Jolly C, Truy E. Radiologic study of a disposable drug delivery intracochlear catheter. *Otol Neurotol*. 2011;32(2):217-22.
- James CJ, Fraysse B, Deguine O, Lenarz T, Mawman D, Ramos A, Ramsden R, Sterkers O. Combined electroacoustic stimulation in conventional candidates for cochlear implantation. *Audiol Neurotol*. 2006;11 Suppl 1:57-62.

- Kiefer J, Tillein J, von Ilberg C, Pfennigdorff T, Stürzebecher E, Klinke R, Gstöttner W. Fundamental aspects and first results of the clinical application of combined electric and acoustic stimulation of the auditory system. *Cochlear Implants – An update*:569-576. Kluger publications, 2002.
- Kiefer J, Gstottner W, Baumgartner W, Pok SM, Tillein J, Ye Q, von Ilberg C. Conservation of low-frequency hearing in cochlear implantation. *Acta Otolaryngol.* 2004;124(3):272-80.
- Kiefer J, Pok M, Adunka O, Stürzebecher E, Baumgartner W, Schmidt M, Tillein J, Ye Q, Gstottner W. Combined electric and acoustic stimulation of the auditory system: results of a clinical study. *Audiol Neurotol.* 2005;10(3):134-44.
- Kong YY, Braida LD. Cross-frequency integration for consonant and vowel identification in bimodal hearing. *J Speech Lang Hear Res.* 2011;54(3):959-80.
- Kong YY, Carlyon RP. Improved speech recognition in noise in simulated binaurally combined acoustic and electric stimulation. *J Acoust Soc Am.* 2007;121(6):3717-27.
- Kong YY, Stickney GS, Zeng FG. Speech and melody recognition in binaurally combined acoustic and electric hearing. *J Acoust Soc Am.* 2005;117(3 Pt 1):1351-61.
- Lee J, Nadol JB Jr, Eddington DK. Depth of electrode insertion and postoperative performance in humans with cochlear implants: a histopathologic study. *Audiol Neurotol.* 2010;15(5):323-31.
- Lenarz T, Stöver T, Buechner A, Lesinski-Schiedat A, Patrick J, Pesch J. Hearing conservation surgery using the Hybrid-L electrode. Results from the first clinical trial at the Medical University of Hannover. *Audiol Neurotol.* 2009;14 Suppl 1:22-31.
- Lenarz T, Stover T, Buechner A, Paasche G, Briggs R, Risi F, Pesch J, Battmer RD. Temporal bone results and hearing preservation with a new straight electrode. *Audiol Neurotol.* 2006;11 Suppl 1:34-41.

- Liberman MC, Dodds LW. Single-neuron labeling and chronic cochlear pathology. III. Stereocilia damage and alterations of threshold tuning curves. *Hear Res.* 1984;16(1):55-74.
- Lina-Granade G, Gallego S, Thai-Van H, Truy E. Appareillage auditif conventionnel par voie aérienne. EMC (Elsevier Masson SAS, Paris), Oto-rhino-laryngologie, 2010;20-185-C-15 : 1-15.
- Loizou PC. Speech processing in vocoder-centric cochlear implants. *Adv Otorhinolaryngol.* 2006;64:109-43.
- Lorenzi C, Gilbert G, Carn H, Garnier S, Moore BC. Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proc Natl Acad Sci U S A.* 2006;103(49):18866-9.
- Luetje CM, Brackman D, Balkany TJ, Maw J, Baker RS, Kelsall D, Backous D, Miyamoto R, Parisier S, Arts A. Phase III clinical trial results with the Vibrant Soundbridge implantable middle ear hearing device: a prospective controlled multicenter study. *Otolaryngol Head Neck Surg.* 2002;126(2):97-107.
- Luntz M, Shpak T, Weiss H. Binaural-bimodal hearing: concomitant use of a unilateral cochlear implant and a contralateral hearing aid. *Acta Otolaryngol.* 2005 Aug;125(8):863-9.
- Mackersie CL, Crocker TL, Davis RA. Limiting high-frequency hearing aid gain in listeners with and without suspected cochlear dead regions. *J Am Acad Audiol.* 2004;15(7):498-507.
- McCreery DB, Agnew WF, Yuen TG, Bullara L. Charge density and charge per phase as cofactors in neural injury induced by electrical stimulation. *IEEE Trans Biomed Eng.* 1990;37(10):996-1001.

- McDermott HJ, Lech M, Kornblum MS, Irvine DR. Loudness perception and frequency discrimination in subjects with steeply sloping hearing loss: possible correlates of neural plasticity. *J Acoust Soc Am*. 1998;104(4):2314-25.
- McDermott H, Sucher C, Simpson A. Electro-acoustic stimulation. Acoustic and electric pitch comparisons. *Audiol Neurotol*. 2009;14 Suppl 1:2-7.
- Middlebrooks JC. Effects of cochlear-implant pulse rate and inter-channel timing on channel interactions and thresholds. *J Acoust Soc Am*. 2004;116:452-68.
- Mok M, Grayden D, Dowell RC, Lawrence D. Speech perception for adults who use hearing aids in conjunction with cochlear implants in opposite ears. *J Speech Lang Hear Res*. 2006;49(2):338-51.
- Moore BCJ. Cochlear hearing loss - Physiological, Psychological and Technical Issues. Second edition, Wiley, 2007.
- Moore BCJ. Dead Regions in the Cochlea: Diagnosis, Perceptual Consequences, and Implications for the Fitting of Hearing Aids. *Trends Amplif*. 2001;5(1):1-34.
- Moore BC, Glasberg BR, Stone MA. New version of the TEN test with calibrations in dB HL. *Ear Hear*. 2004;25(5):478-87.
- Moore BC, Huss M, Vickers DA, Glasberg BR, Alcántara JI. A test for the diagnosis of dead regions in the cochlea. *Br J Audiol*. 2000;34(4):205-24.
- Moore BC, Vinay SN. Enhanced discrimination of low-frequency sounds for subjects with high-frequency dead regions. *Brain*. 2009;132(Pt 2):524-36.
- Nadol JB Jr, Ketten DR, Burgess BJ. Otopathology in a case of multichannel cochlear implantation. *Laryngoscope*. 1994;104(3 Pt 1):299-303.

- Nadol JB Jr, Shiao JY, Burgess BJ, Ketten DR, Eddington DK, Gantz BJ, Kos I, Montandon P, Coker NJ, Roland JT Jr, Shallop JK. Histopathology of cochlear implants in humans. *Ann Otol Rhinol Laryngol*. 2001;110(9):883-91.
- Nadol JB Jr, Young YS, Glynn RJ. Survival of spiral ganglion cells in profound sensorineural hearing loss: implications for cochlear implantation. *Ann Otol Rhinol Laryngol*. 1989;98(6):411-6.
- Ni D, Shepherd RK, Seldon HL, Xu SA, Clark GM, Millard RE. Cochlear pathology following chronic electrical stimulation of the auditory nerve. I: Normal hearing kittens. *Hear Res*. 1992;62(1):63-81.
- Novak MA, Black JM, Koch DB. Standard cochlear implantation of adults with residual low-frequency hearing: implications for combined electro-acoustic stimulation. *Otol Neurotol*. 2007;28(5):609-14.
- Olson AD, Shinn JB. A systematic review to determine the effectiveness of using amplification in conjunction with cochlear implantation. *J Am Acad Audiol*. 2008;19(9):657-71.
- Oxenham AJ. Pitch perception and auditory stream segregation: implications for hearing loss and cochlear implants. *Trends Amplif*. 2008;12(4):316-31.
- Paasche G, Gibson P, Averbeck T, Becker H, Lenarz T, Stöver T. Technical report: modification of a cochlear implant electrode for drug delivery to the inner ear. *Otol Neurotol*. 2003;24(2):222-7.
- Patrick JF, Busby PA, Gibson PJ. The development of the Nucleus Freedom Cochlear implant system. *Trends Amplif*. 2006 Dec;10(4):175-200.
- Prentiss S, Sykes K, Staecker H. Partial deafness cochlear implantation at the University of Kansas: techniques and outcomes. *J Am Acad Audiol*. 2010;21(3):197-203.

- Qin MK, Oxenham AJ. Effects of introducing unprocessed low-frequency information on the reception of envelope-vocoder processed speech. *J Acoust Soc Am.* 2006;119(4):2417-26.
- Rankovic CM. Articulation index predictions for hearing-impaired listeners with and without cochlear dead regions. *J Acoust Soc Am.* 2002;111(6):2545-8.
- Reiss LA, Turner CW, Erenberg SR, Gantz BJ. Changes in pitch with a cochlear implant over time. *J Assoc Res Otolaryngol.* 2007;8(2):241-57.
- Rhodes RC. Discrimination of filtered CNC lists by normals and hypacusics. *J. Aud. Res.* 1966;6:129-133
- Richardson RT, Noushi F, O'Leary S. Inner ear therapy for neural preservation. *Audiol Neurotol.* 2006;11(6):343-56.
- Rosen S. Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos Trans R Soc Lond B Biol Sci.* 1992 29;336(1278):367-73.
- Rosen S, Faulkner A, Wilkinson L. Adaptation by normal listeners to upward spectral shifts of speech: implications for cochlear implants. *J Acoust Soc Am.* 1999;106(6):3629-36.
- Rubinstein JT, Wilson BS, Finley CC, Abbas PJ. Pseudospontaneous activity: stochastic independence of auditory nerve fibers with electrical stimulation. *Hear Res.* 1999;127:108-18.
- Sagi E, Meyer TA, Kaiser AR, Teoh SW, Svirsky MA. A mathematical model of vowel identification by users of cochlear implants. *J Acoust Soc Am.* 2010;127(2):1069-83.
- Shannon RV. A model of safe levels for electrical stimulation. *IEEE Trans Biomed Eng.* 1992;39(4):424-6.
- Shannon RV. Forward masking in patients with cochlear implants. *J Acoust Soc Am.* 1990;88:741-4.

- Shepherd RK, Clark GM, Black RC. Chronic electrical stimulation of the auditory nerve in cats. Physiological and histopathological results. *Acta Otolaryngol Suppl.* 1983a;399:19-31.
- Shepherd R K, Clark GM, Black RC, Patrick JF. The histopathological effects of chronic electrical stimulation of the cat cochlea. *Journal of Laryngology and Otology* 1983b;97:333-341.
- Shepherd RK, Hatsushika S, Clark GM. Electrical stimulation of the auditory nerve: the effect of electrode position on neural excitation. *Hear Res.* 1993;66(1):108-20.
- Simpson A, McDermott HJ, Dowell RC. Benefits of audibility for listeners with severe high-frequency hearing loss. *Hear Res.* 2005;210(1-2):42-52.
- Skarżyński H, Lorens A, D'Haese P, Walkowiak A, Piotrowska A, Sliwa L, Anderson I. Preservation of residual hearing in children and post-lingually deafened adults after cochlear implantation: an initial study. *ORL J Otorhinolaryngol Relat Spec.* 2002;64(4):247-53.
- Skarzynski H, Lorens A, Piotrowska A, Anderson I. Partial deafness cochlear implantation provides benefit to a new population of individuals with hearing loss. *Acta Otolaryngol.* 2006;126(9):934-40.
- Skinner MW, Ketten DR, Holden LK, Harding GW, Smith PG, Gates GA, Neely JG, Kletzker GR, Brunsden B, Blocker B. CT-derived estimation of cochlear morphology and electrode array position in relation to word recognition in Nucleus-22 recipients. *J Assoc Res Otolaryngol.* 2002;3(3):332-50.
- Spelman FA. Cochlear electrode arrays: past, present and future. *Audiol Neurotol.* 2006;11(2):77-85.
- Spoendlin H, Schrott A. The spiral ganglion and the innervation of the human organ of Corti. *Acta Otolaryngol.* 1988;105(5-6):403-10.

- Stakhovskaya O, Sridhar D, Bonham BH, Leake PA. Frequency map for the human cochlear spiral ganglion: implications for cochlear implants. *J Assoc Res Otolaryngol.* 2007;8(2):220-33.
- Summers V. Do tests for cochlear dead regions provide important information for fitting hearing aids? *J Acoust Soc Am.* 2004;115(4):1420-3.
- Thai-Van H, Micheyl C, Moore BC, Collet L. Enhanced frequency discrimination near the hearing loss cut-off: a consequence of central auditory plasticity induced by cochlear damage? *Brain.* 2003;126(Pt 10):2235-45.
- Thai-Van H, Micheyl C, Norena A, Collet L. Local improvement in auditory frequency discrimination is associated with hearing-loss slope in subjects with cochlear damage. *Brain.* 2002;125(Pt 3):524-37.
- Todt I, Seidl RO, Gross M, Ernst A. Comparison of different vibrant soundbridge audioprocessors with conventional hearing AIDS. *Otol Neurotol.* 2002;23(5):669-73.
- Truy E, Bouccara D, Sterkers O, Triglia JM. Techniques chirurgicales d'implantation d'aides auditives en otoneurologie. EMC (Elsevier Masson SAS, Paris), Techniques chirurgicales – Tête et cou, 2009 ;46-038:1-20.
- Truy E, Philibert B, Vesson JF, Labassi S, Collet L. Vibrant soundbridge versus conventional hearing aid in sensorineural high-frequency hearing loss: a prospective study. *Otol Neurotol.* 2008;29(5):684-7.
- Turner CW, Cummings KJ. Speech audibility for listeners with high-frequency hearing loss. *Am J Audiol.* 1999;8(1):47-56.
- Turner CW, Gantz BJ, Vidal C, Behrens A, Henry BA. Speech recognition in noise for cochlear implant listeners: benefits of residual acoustic hearing. *J Acoust Soc Am.* 2004;115(4):1729-35.

- Turner CW, Reiss LA, Gantz BJ. Combined acoustic and electric hearing: preserving residual acoustic hearing. *Hear Res.* 2008 Aug;242(1-2):164-71.
- Tykocinski, M., Shepherd, R.K., Clark, G.M. Reduction in excitability of the auditory nerve following electrical stimulation at high stimulus rates. *Hear Res.* 1995;88:124-42.
- Tyler RS, Parkinson AJ, Wilson BS, Witt S, Preece JP, Noble W. Patients utilizing a hearing aid and a cochlear implant: speech perception and localization. *Ear Hear.* 2002;23(2):98-105.
- Tyler RS, Parkinson AJ, Woodworth GG, Lowder MW, Gantz BJ. Performance over time of adult patients using the Ineraid or nucleus cochlear implant. *J Acoust Soc Am.* 1997;102(1):508-22.
- Tysome JR, Moorthy R, Lee A, Jiang D, O'Connor AF. Systematic review of middle ear implants: do they improve hearing as much as conventional hearing AIDS? *Otol Neurotol.* 2010;31(9):1369-75.
- Usami S, Moteki H, Suzuki N, Fukuoka H, Miyagawa M, Nishio SY, Takumi Y, Iwasaki S, Jolly C. Achievement of hearing preservation in the presence of an electrode covering the residual hearing region. *Acta Otolaryngol.* 2011;131(4):405-12.
- Uziel A, Mondain M, Hagen P, Dejean F, Doucet G. Rehabilitation for high-frequency sensorineural hearing impairment in adults with the symphonix vibrant soundbridge: a comparative study. *Otol Neurotol.* 2003;24(5):775-83.
- Van de Heyning P, Punte AK. Electric acoustic stimulation: a new era in prosthetic hearing rehabilitation. *Adv Otorhinolaryngol.* 2010;67:1-5.
- Vermeire K, Nobbe A, Schleich P, Nopp P, Voormolen MH, Van de Heyning PH. Neural tonotopy in cochlear implants: an evaluation in unilateral cochlear implant patients with unilateral deafness and tinnitus. *Hear Res.* 2008;245(1-2):98-106.

- Vestergaard MD. Benefit from amplification of high frequencies in hearing impaired: aspects of cochlear dead regions and auditory acclimatization. Technical University of Denmark, Ørsted-DTU Acoustic Technology, DK-2800 Lyngby, 2004.
- Vestergaard MD. Dead regions in the cochlea: implications for speech recognition and applicability of articulation index theory. *Int J Audiol*. 2003;42(5):249-61.
- Vickers DA, Moore BC, Baer T. Effects of low-pass filtering on the intelligibility of speech in quiet for people with and without dead regions at high frequencies. *J Acoust Soc Am*. 2001;110(2):1164-75.
- von Ilberg C, Kiefer J, Tillein J, Pfenningdorff T, Hartmann R, Stürzebecher E, Klinke R. Electric-acoustic stimulation of the auditory system. New technology for severe hearing loss. *ORL J Otorhinolaryngol Relat Spec*. 1999;61(6):334-40.
- Wilson, B.S. 2011. Low pulse rate cochlear implant stimulation in conjunction with a separate representation of fundamental frequencies and voiced/unvoiced distinctions. US patent US 2011/0066210 A1.
- Wilson BS, Dorman MF. Cochlear implants: a remarkable past and a brilliant future. *Hear Res*. 2008;242(1-2):3-21.
- Wilson BS, Finley CC, Lawson DT, Wolford RD, Eddington DK, Rabinowitz WM. Better speech recognition with cochlear implants. *Nature*. 1991;352(6332):236-8.
- Woodson EA, Reiss LA, Turner CW, Gfeller K, Gantz BJ. The Hybrid cochlear implant: a review. *Adv Otorhinolaryngol*. 2010;67:125-34.
- Xi X, Ji F, Han D, Hong M, Chen A. Electrode interaction in cochlear implant recipients: comparison of straight and contour electrode arrays. *ORL J Otorhinolaryngol Relat Spec*. 2009;71(4):228-37.
- Xu SA, McAnally KI, Xu J, Clark GM. Comparison of half-band and full-band electrodes for intracochlear electrical stimulation. *Ann Otol Rhinol Laryngol*. 1993;102(5):363-7.

- Xu J, Shepherd RK, Millard RE, Clark GM. Chronic electrical stimulation of the auditory nerve at high stimulus rates: a physiological and histopathological study. *Hear Res.* 1997;105(1-2):1-29.
- Yates GK. Basilar membrane nonlinearity and its influence on auditory nerve rate-intensity functions. *Hear Res.* 1990;50(1-2):145-62.
- Zeng FG, Rebscher S, Harrison WV, Sun X, Feng H. Cochlear Implants: System Design, Integration and Evaluation. *IEEE Rev Biomed Eng.* 2008 1;1:115-142.
- Zhang T, Dorman MF, Spahr AJ. Information from the voice fundamental frequency (F0) region accounts for the majority of the benefit when acoustic stimulation is added to electric stimulation. *Ear Hear.* 2010;31(1):63-9.
- Zwicker E, Schorn K. Psychoacoustical tuning curves in audiology. *Audiology.* 1978;17(2):120-40.

SPÉCIFICITÉS DE L'IMPLANT ÉLECTRO-ACOUSTIQUE : INDICATIONS, INTERFACE BIOÉLECTRIQUE ET STRATÉGIE DE CODAGE

Le clinicien se trouve parfois confronté à des sujets qui présentent une surdité supérieure à 90 dB HL au-delà de 1 kHz avec une audition résiduelle dans les fréquences graves. Pour réhabiliter les hautes fréquences, il existe aujourd'hui différentes technologies : amplification conventionnelle, compression fréquentielle, implant cochléaire et depuis une dizaine d'année la stimulation électro-acoustique EAS qui consiste à stimuler acoustiquement les sons graves et électriquement les sons aigus via un implant cochléaire.

La première partie de cette thèse a consisté à identifier les facteurs qui influencent les capacités des patients sourds partiels à traiter l'information basse fréquence de la parole. Nous avons utilisé un test d'audiométrie vocale filtrée passe-bas. Nos résultats indiquent que les scores d'intelligibilité de la parole sont positivement corrélés avec la durée de la surdité. Ceci signifie qu'avec le temps, ces sujets malentendants apprennent à comprendre avec cette audition type filtre passe-bas, à tel point que certains ont des performances supra-normales pour l'utilisation des basses fréquences. Nos résultats montrent également une corrélation négative entre l'âge d'apparition de la surdité et les scores d'intelligibilité. Ce test pourra aider le clinicien à mieux cibler l'appareillage le plus adapté à chaque profil de patient.

La seconde partie de cette thèse, consacrée à l'EAS, a consisté à évaluer par des simulations chez le normo-entendant, diverses stratégies de codage du son par l'implant EAS. Actuellement, la stratégie utilisée pour l'EAS est calquée sur celle de l'implant cochléaire et nos résultats suggèrent que cette stratégie peut être optimisée.

SPECIFICITIES OF ELECTRIC-ACOUSTIC STIMULATION : INDICATIONS, BIOELECTRICAL INTERFACE AND CODING STRATEGY

Clinicians may face patients who have a deafness superior to 90 dB HL above 1 kHz with good low-frequency residual hearing. Today, several technologies are available to provide high frequencies: conventional amplification, frequency compression, cochlear implant since about 10 years Electric-Acoustic Stimulation EAS which consists in stimulating acoustically low frequencies while stimulating electrically high frequency sounds via a cochlear implant.

The first part of this dissertation consisted in identifying the factors which may influence abilities of partially deaf subjects to process low-frequency speech information. We used a low-pass filtered speech test. Our results show that speech intelligibility scores are positively correlated to the duration of deafness. This means that these hearing-impaired subjects learn to understand with this lowpass-like hearing, in such a way that some of them exhibit supranormal abilities for the processing of low-frequency sounds. Our results also show a negative correlation between the age at onset of deafness and speech intelligibility scores. This test may help the clinician to better evaluate which device would be best for every patient's profile.

The second part of this dissertation, about EAS, consisted in evaluating through simulations in normal-hearing listeners, several coding strategies by the EAS implant. Now the strategy used for EAS duplicates the strategy used by cochlear implants and our results suggest that this strategy could be optimized.

MOTS-CLES : Implantation cochléaire ; Stimulation Electrique-Acoustique ; Surdité partielle ; Parole filtrée passe-bas ; Intelligibilité de la parole ; Durée de privation ; Age d'apparition de la surdité ; Combinaison des informations ; Stratégie de codage ; Vocoder.